THEORETICAL AND NUMERICAL ASPECTS OF COALESCING OF EIGENVALUES AND SINGULAR VALUES OF PARAMETER DEPENDENT MATRICES

A Thesis
Presented to
The Academic Faculty

by

Alessandro Pugliese

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Mathematics

Georgia Institute of Technology
August 2008

THEORETICAL AND NUMERICAL ASPECTS OF COALESCING OF EIGENVALUES AND SINGULAR VALUES OF PARAMETER DEPENDENT MATRICES

Approved by:

Professor Luca Dieci, Advisor
School of Mathematics
Georgia Institute of Technology

Professor Shui-Nee Chow
School of Mathematics
Georgia Institute of Technology

Professor Michael Loss
School of Mathematics
Georgia Institute of Technology

Professor Yingjie Liu
School of Mathematics
Georgia Institute of Technology

Professor Erik Verriest
School of Electrical and Computer Engineering
Georgia Institute of Technology

Date Approved: April 22, 2008

Ad Agnese,

la mia forza.

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor Luca Dieci for his guidance, support and friendship. I consider myself very fortunate to have met him along my way.

I would like to thank Professors Shui-Nee Chow, Yingjie Liu, Michael Loss and Erik Verriest for serving on my defense committee.

I acknowledge financial support from several sources: The School of Mathematics at Georgia Tech, the National Science Foundation (grant DMS 0139895), the numerical analysis group at the *Universitá di Bari* and the D'Onofrio fellowship.

Thanks to Luciano Lopez for having believed in me, thanks to my parents and my sister for their unconditional love and enthusiastic support, and thanks to all the friends I met in Atlanta for the wonderful times we had together.

Dulcis in fundo, thanks to Agnese, to whom this effort is dedicated.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

In this thesis, we study the problem of coalescing singular values of smooth real-valued parameter-dependent matrices, or of coalescing eigenvalues of smooth real-valued parameter-dependent symmetric matrices. In particular, we are interested in detecting, and accurately approximating, parameters' values where the coalescing occurs. Our numerical techniques are based on theoretical results which are not local nor perturbative, but rather global in nature.

We begin with a review of known results that concern the smoothness of decompositions of matrix functions and some background material. Then, we state and prove several theoretical results that concern smooth matrix functions of two parameters, starting with the simplest (2×2) case and ending with the most general $(n \times n)$ case. In all these results, the existence of points where the singular values of a smooth matrix function A coalesce in a bounded simply connected domain Ω is related to the periodic structure of the smooth singular values decomposition (SVD) of A computed around the boundary of Ω. This interplay between *coalescing points* and periodicity of a smooth SVD forms the backing of algorithms for the detection and approximation of coalescing points in planar regions. We present these algorithms in detail, and show performance of their implementation through several examples. Finally, we present numerical techniques for the continuation of curves of coalescing singular values of real-valued matrices that depend on three parameters, and illustrate, with two examples, how these techniques can be used to tackle the case of coalescing singular values for complex-valued matrices that depend on three parameters.

CHAPTER I

INTRODUCTION AND BACKGROUND

1.1 Motivation and Scope

The study of variation of eigenvalues (and eigenvectors) of a matrix depending on parameters, in function of the parameters, is a classical problem and has been extensively studied. Many are its applications throughout science and engineering. Notably, it appears in the study of dynamical systems, where stability and bifurcation studies of fixed points and/or periodic orbits typically reduce to studying spectral properties of parameter dependent matrices (Jacobians or monodromy matrices).

In this thesis, we will be concerned chiefly with the problem of coalescing singular values for general (i.e. non-symmetric) real matrix functions, or, almost equivalently, of coalescing eigenvalues for real symmetric matrix functions. Our main focus will be on finding coalescing points for singular values (and eigenvalues) of matrix functions that depend smoothly on two parameters.

The problem of finding points where eigenvalues coalesce has received an incredible amount of attention in several applied sciences, such as Quantum Mechanics, Chemical Physics, Molecular Physics and Quantum Chemistry, just to name a few. We will further discuss some of these applications in Section 1.3. On the contrary, this problem doesn't appear to have received the same deal of attention in the mathematics community.

In this thesis we accomplish the following tasks. First, in the remainder of this introduction, we review some fundamental existing results that concern matrices depending on parameters. Second, we present our contribution to the subject, both in terms of theoretical results and numerical algorithms for real matrix functions that

depend on two parameters. This will be the content of Chapters 2, 3 and 4. Finally, in Chapter 5, we give extensions to the case of real, and complex, matrix functions that depend on three parameters.

The content of Chapters 1 through 4 is based on our papers [9] and [10].

1.2 Definition of the Problem

For a matrix $A \in \mathbb{R}^{n\times n}$, the Singular Value Decomposition (SVD) is the decomposition $A = U\Sigma V^T$, where $U, V \in \mathbb{R}^{n\times n}$ are orthogonal and Σ is diagonal, with diagonal given by the singular values $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. The SVD is a well known, and widely used, tool in Linear Algebra, as it reveals a lot of information associated to A. For instance, orthonormal bases for the range and the null space of A, as well as for the range and null space of A^T, can all be obtained easily from the factors of the SVD; see [13].

If the matrix A depends smoothly on parameters, it is natural to inquire into the differentiability of the factors of the SVD of A with respect to those parameters. Unfortunately, except for the one-parameter case, the factors in the SVD of A in general do not vary smoothly when singular values of A coalesce. See Section 1.4 for more details and references.

Moreover, when singular values of the function A, say the k-th and $(k+1)$-st, coalesce, there is an immediate lack of uniqueness in the matrix of best approximation of rank k to A. In this sense, coalescing of singular values represents (a local) bifurcation phenomenon. And, similarly to what one does when studying bifurcation phenomena, it is thus important to locate parameter values where singular values coalesce: *coalescing points*. This is the problem that we consider in this thesis: To accurately locate coalescing points.

For reasons that will soon be clear (see Section 1.2.1), we will mainly be interested in matrices smoothly depending on two parameters, and taking values in $\mathbb{R}^{n\times n}$; only

Chapter 5 will be devoted to matrix functions depending on three parameters (both real and complex valued).

The problem that will be considered in this thesis can be stated as follows. In what follows, we may think of a smooth function as one which is at least $\mathcal{C}^1$; we will be more precise soon.

Problem 1. Given a smooth matrix function $A : \mathbb{R}^2 \to \mathbb{R}^{n\times n}$, find values $x \in \mathbb{R}^2$ where the singular values of A coalesce.

The SVD of a matrix A is intimately related to several symmetric eigenproblems, such as those for A^TA and AA^T. Therefore, it will be natural to consider also the following problem, nearly equivalent to Problem 1:

Problem 2. Given a smooth matrix function $A : \mathbb{R}^2 \to \mathbb{R}^{n\times n}$, symmetric, find values $x \in \mathbb{R}^2$ where the eigenvalues of A coalesce.

1.2.1 Genericity and Codimension of Coalescing Singular Values

A very important issue to address, when dealing with functions at large, is that of genericity. Loosely speaking, a generic property is a property that one expects to hold for a "typical function" (see [17] for background on this concept). Addressing this issue helps distinguishing between what is a pathological –and somewhat artificial– behavior from what is a "regular" behavior and should be expected. This is of primary importance, both from the theoretical and numerical point of view.

In 1929, J. von Neumann and E. Wigner gave a fundamental result (see [32]) which is directly related to the genericity of having a pair of coalescing singular values. Their result, known as *non-crossing rule*, states that having a pair of coalescing eigenvalues for a real symmetric matrix valued function, and hence a pair of coalescing singular values for a general (i.e. non-symmetric) matrix valued function, is a codimension-2 phenomenon. Their proof of this fact is quite simple, and will be hinted at in Section

2.1[1].

For our purposes, the main consequence of this fact is that having a pair of coalescing singular values is a generic property for smooth matrix valued functions depending on 2 (or more) parameters, but not for functions depending on 1 parameter. More precisely, we should not expect a smooth matrix valued function depending on 1 parameter to have coalescing singular values, and in case a pair of singular values do coalesce, the coalescing should break down under almost any arbitrarily small perturbation (we will come back to this in Chapter 3). On the other hand, for two-parameter matrix functions, the generic statement is that coalescing of singular values must be expected, that it should occur at isolated points in parameter space, and that it should persist under small perturbations.

This is the reason why it is natural to focus on the class of two-parameter smooth matrix functions when addressing Problems 1 and 2.

Remark 1.2.1. The issue of genericity of having multiple eigenvalues for a smooth real symmetric matrix function (or singular values for a non-symmetric function) has been addressed also in [7]. The argument therein, based on the use of the Whitney topology (again, see [16]), can be extended to the case of more complicated coalescing situations. In particular, it can be used to prove that coalescing of three (consecutive) eigenvalues of a smooth symmetric function, at the same parameter value, is a codimension 5 phenomenon, while coalescing of two pairs of eigenvalues at the same parameter value is a codimension 4 phenomenon. Therefore, such phenomena are highly non-generic for two-parameter matrix functions. This justifies one of the hypotheses that we will make for our results in Chapter 2: We will assume that eigenvalues (or singular values) coalesce "one pair at a time", as it is generically expected.

[1] see Remark 2.1.10

Remark 1.2.2. We need to stress one more point: Coalescing of singular values is a codimension 2 phenomenon for *non-symmetric* matrix functions. For real symmetric matrices, it has codimension 1. We will justify this in Chapter 2 (see Remark 2.1.17). But, of course, real symmetric functions possess extra structure, whereas throughout this thesis we will always consider the problem of coalescing singular values within the class of matrix functions that do not posses any additional particular structure (like symmetry), as that may lower the codimension of the phenomenon.

1.3 Applications in Science and Engineering

The phenomenon of coalescing of eigenvalues plays an important role in several areas of Physics and Chemistry, such as Quantum Mechanics and Quantum Chemistry. Therein, eigenvalues represent energies of atomic arrangements. In the Chemistry literature, points where eigenvalues of parameter-dependent hermitian matrices coalesce are typically referred to as points of *conical intersection* (see [33]); in Quantum Physics, they have been called *diabolical points* (see [4]). Indeed, in those areas, the presence of conical intersections signifies the breakdown of the so-called *Born-Oppenheimer approximation* and the concurrent increased importance of *nonadiabatic coupling terms*. The earliest reference we are aware of is the previously cited seminal work of J. Von Neumann and E. Wigner, [32]. Later, the existence of conical intersections has been linked to interesting phenomena such as the Longuet-Higgins effect (see [15]) and Berry's geometric phase effect (see [3]). Recently, the book by Baer ([2]), together with the review article by Yarcony ([33]), have provided an account of the diminishing impact of the Born-Oppenheimer approximation, as well as of the consequent importance of conical intersections within nonadiabatic processes. A search for the words "conical intersection" within the *J. of Chemical Physics* returns 321 hits over the last 40 years, which testifies on the great deal of attention that the Chemical Physics community has devoted to this subject. By comparison,

the same search performed within the *MathSciNet* database returns no hits in any mathematics journal.

1.3.1 Why the Name "Conical Intersection"

Singular values and eigenvalues of a two-parameter dependent matrix function describe surfaces in $\mathbb{R}^3$. Those surfaces are generally referred to as *potential energy surfaces* within the adiabatic or Born-Oppenheimer approximation. In the Example 1.3.1 below, we justify the name of *conical* intersections, which is given to the points where those surfaces coalesce. We elucidate only the case of a symmetric function (similar considerations apply for coalescing singular values of a general function).

Example 1.3.1. Consider the symmetric matrix valued function below

$$P(x) \;=\; \begin{bmatrix} a(x) & b(x) \\ b(x) & d(x) \end{bmatrix} \;, \quad x = (x_1, x_2) \in \Omega \,,$$

and let ξ_0 be a point where the eigenvalues of P coalesce. In other words, since the (continuous) eigenvalues have the form

$$\lambda_{\pm}(x) \;=\; \frac{a(x)+d(x)}{2} \pm \frac{1}{2}\sqrt{(a(x)-d(x))^2 + 4b(x)^2}\,,$$

at ξ_0 we have $(a-d)(\xi_0) = 0$ and $b(\xi_0) = 0$. Assume that $\nabla(a-d)(\xi_0) \neq 0$ and $\nabla b(\xi_0) \neq 0$ and that the Jacobian $\begin{bmatrix} \nabla(a-d) \\ \nabla b \end{bmatrix}_{\xi_0}$ is nonsingular (this will be later defined as *transversal intersection* at ξ_0).

We write the eigenvalues in the form $\lambda_{\pm} = \frac{(a+d)(x)}{2} \pm \frac{1}{2}\sqrt{h(x)}$ and expand the function $h(x)$ at ξ_0. We get

$$h(x) = h(\xi_0) + \nabla h(\xi_0)(x-\xi_0) + \frac{1}{2}(x-\xi_0)^T H(\xi_0)(x-\xi_0) + \dots \,,$$

and a simple computation gives $h(\xi_0) = 0$, $\nabla h(\xi_0) = 0$, and

$$H(\xi_0) = 2 \begin{bmatrix} [(a-d)_{x_1}]^2 + 4(b_{x_1})^2 & (a-d)_{x_1}(a-d)_{x_2} + 4b_{x_1}b_{x_2} \\ (a-d)_{x_1}(a-d)_{x_2} + 4b_{x_1}b_{x_2} & [(a-d)_{x_2}]^2 + 4(b_{x_2})^2 \end{bmatrix}_{\xi_0} .$$

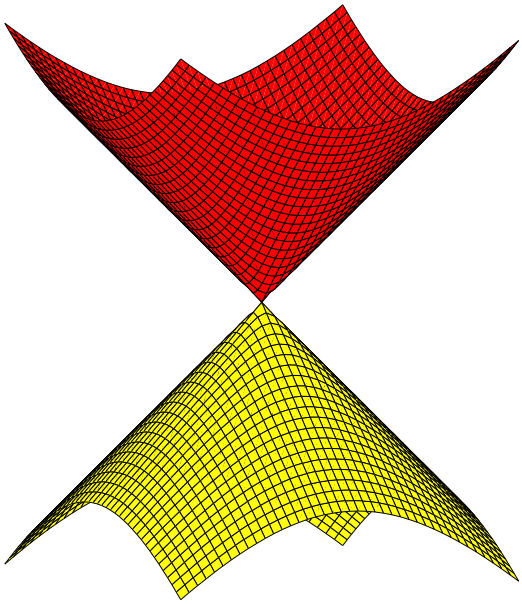

Figure 1: An ideal conical intersection.

Therefore, $\det H(\xi_0) = 16\big[b_{x_1}(a-d)_{x_2} - b_{x_2}(a-d)_{x_1}\big]^2(\xi_0)$ and thus $\det H(\xi_0) = 0$ if and only if at ξ_0:

$$\begin{bmatrix}(a-d)_{x_1} & (a-d)_{x_2}\end{bmatrix}\begin{bmatrix}0 & 1\\ -1 & 0\end{bmatrix}\begin{bmatrix}b_{x_1}\\ b_{x_2}\end{bmatrix} \;=\; 0\,,$$

that is $\nabla(a-d)(\xi_0)$ is parallel to $\nabla b(\xi_0)$, and thus the Jacobian $\begin{bmatrix}\nabla(a-d)\\ \nabla b\end{bmatrix}_{\xi_0}$ will be singular, which we have excluded. We can conclude that $H(\xi_0)$ is positive definite, and that in the vicinity of ξ_0 the eigenvalues have the form $\lambda_\pm = \frac{(a+d)(x)}{2} \pm \frac{1}{2}\sqrt{\|z\|^2 + O(\|x-\xi_0\|^4)}$, where $z = H^{1/2}(\xi_0)\ (x-\xi_0)$. As a consequence, the eigenvalues behave like $\pm$ the norm $\|z\|$, which is a double cone (see Figure 1).

1.3.2 Conical Intersections and "Avoided Crossings" in Structural Dynamics

The dependence of eigenvalues (and eigenvectors) upon system parameters is of interest also in Engineering, for example in Structural Dynamics. In this section, we elucidate this application in some detail.

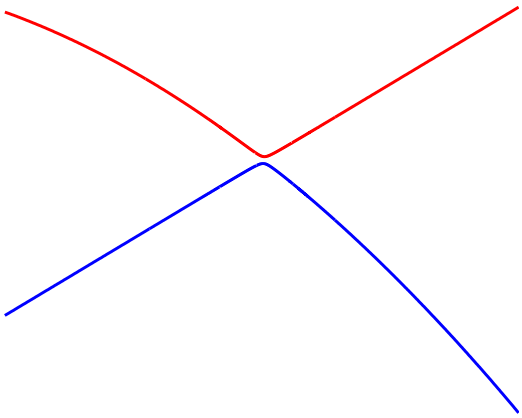

Figure 2: Avoided crossing between two eigenvalues' curves.

The study of vibrations of a mechanical system begins by considering the so-called free-vibration problem. This problem is associated to an eigenvalue problem, usually symmetric, where eigenvalues represent frequency of vibrations of the mechanical structure, and eigenvectors (eigenfunctions) represent the corresponding modes of vibration. These free vibration modes occur when the mechanical structure is allowed to vibrate freely, with no externals forces acting on it. The corresponding frequencies of vibration are called the *natural* frequencies of the structure. When external forces act on the structure, they induce a response of the structure that depends on how the frequency of the external force and the natural frequencies of the structure are related (e.g., resonant behavior when the forcing frequency is near the natural frequency). In this context, coalescing of eigenvalues gives rise to an interesting phenomenon: Two different modes of vibration are induced by the excitation of one single natural frequency.

Perhaps surprisingly, in engineering applications it is more frequent to encounter the phenomenon of *veering* (of eigenvalues' curves), also known as *avoided crossing*, see Figure 2: Instead of coalescing, eigenvalues come very close to coalesce, but suddenly "veer away". Avoided crossings are typical also of other applied sciences, such as Molecular Physics. The scenario in Figure 2 is ubiquitous in all situations where a single system parameter is varied (true crossings require the variation of two

parameters, as pointed out in Section 1.2.1). We will elucidate this point in Chapter 3.

For some time, in the Engineering literature the phenomenon has not been universally accepted. Leissa (in [20]) demonstrated that it can be an aberration induced by discretization and raised doubt on the validity of approximate solutions (obtained via finite elements methods) that showed the occurrence of veering. On the other hand, Perkins and Mote (in [25]) proved that eigenvalues' curve veering also exists in continuous models.

As it will be observed in Chapter 3, in the veering region the eigenfunctions undergo dramatic changes. More precisely, the eigenfunctions associated to coalescing eigenvalues are interchanged during veering, in a rapid but continuous way. Since different eigenfunctions are associated to different modes of vibration, veering has been associated to hypersensitivity of the structure in its response to external forces: Very small variations in a system parameter and/or in the frequency of the external force induce large differences in the response of the structure (given by the excitation of different modes of vibration). This is also known as "strong mode localization", see [21] and [24].

In Engineering applications, problems usually depend on several parameters. We believe that observing a veering, while one system parameter is varied, may disclose the presence of a conical intersection near the veering point for a problem that depends on two (or more) parameters. In a certain way, Figure 2 can be thought as a slice of Figure 1 near the conical intersection point. With this in mind, by developing algorithms which locate the set of conical intersection points, we will effectively predict where the structure will exhibit the above mentioned hypersensitive behavior.

1.4 Previous Work

In this section, we review some fundamental results concerning matrix functions that depend on one or more parameters, that will be extensively exploited throughout this thesis. First, we introduce some notation.

Notation 1.4.1. With $\Omega \subset \mathbb{R}^2$ we indicate an open simply connected planar region, and $x = (x_1, x_2)$ will be coordinates in Ω. At times we will need to restrict to closed rectangular regions in $\mathbb{R}^2$, which will be indicated by R: $R = \{(x_1, x_2) \in \mathbb{R}^2 : a \leq x_1 \leq b \ , \ c \leq x_2 \leq d \ \}$ (and always $a < b, \ c < d$). We write $A \in \mathcal{C}^e(\Omega, \ \mathbb{R}^{n\times n})$ to indicate a $\mathcal{C}^e$ matrix valued function; typically $e \geq 1$ and finite, but also the analytic case of $A \in \mathcal{C}^\omega$ (A is an analytic function) is of interest. We write $A \in \mathcal{C}^e_\tau(\mathbb{R}, \mathbb{R}^{n\times n})$ to indicate a $\mathcal{C}^e$ and τ-periodic matrix valued function of the real variable t; τ is always assumed to be the minimal positive period. The n singular values of a $(n \times n)$ matrix M will be denoted as $\sigma_1(M) \geq \sigma_2(M) \geq \ldots \geq \sigma_n(M) \geq 0$. Whenever there is no ambiguity, they will be denoted simply as $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$. Analogously, the n eigenvalues of a $(n \times n)$ matrix M will be denoted as $\lambda_1(M), \lambda_2(M), \ldots, \lambda_n(M)$, or simply as $\lambda_1, \lambda_2, \ldots, \lambda_n$.

Several existing works on smooth decomposition of matrix valued functions are concerned with the case of $A(t)$, $t \in [a, b] \subseteq \mathbb{R}$. This case is well understood, and results exist both for the analytic and smooth case. For example, it has been known for a long time (see [18]) that symmetric and analytic matrix valued functions admit analytic Schur decompositions; similarly, analytic matrix valued functions (of one real variable) admit analytic singular value decompositions, SVD (see [5]). Results on smoothness of the factors in the case of $\mathcal{C}^e$, $e \geq 0$, functions of one real variable are given in [7].

Remark 1.4.2. We recall that, in the standard linear algebra setting (again, see [13]), the singular values σ_i, $i = 1, \ldots, n$, of a matrix $A \in \mathbb{R}^{n\times n}$ are ordered so

that $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n \geq 0$. However, in the previously cited works concerned with smoothness of the factors, in order to retain smooth (or analytic) factors, the singular values must be allowed to change ordering when they coalesce and to change sign when they become 0. More properly, one obtains a *signed smooth SVD*. Throughout this thesis, unless otherwise stated, when we talk about the SVD we will always mean the one with ordered and non-negative singular values.

Another problem which has received some attention, for 1-parameter functions, is concerned with functions which are periodic in the parameter; see [31] for early work, and [6] for more recent work related to the symmetric eigendecomposition and SVD. For example, in [6] it was proved that "If the singular values of a function $A \in \mathcal{C}^e_1(\mathbb{R}, \mathbb{R}^{n\times n})$, $e \geq 0$, remain distinct over one period, then A admits a $\mathcal{C}^e$ SVD, where the (signed) singular values are 1-periodic, while the orthogonal functions U and V are either 1-periodic or 2-periodic". There was no indication in [6], and see also [31], of when these functions effectively had period 1 or 2. Perhaps surprisingly, we will realize in this thesis that this fact can be understood by studying functions of 2 parameters: More precisely, by thinking of the 1-parameter periodic function $A \in \mathcal{C}^e_1(\mathbb{R}, \mathbb{R}^{n\times n})$ (with distinct singular values) as a function on a closed loop in two-parameter space, one will have 1-periodic factors U, V, if the singular values of the underlying two-parameter function do not coalesce inside the loop. [In fact, more is true: Depending on whether and how singular values coalesce inside the loop, the factors will have period 1 or 2; see Theorem 2.2.6].

Figure 1 already suggests that, as far as smoothness is concerned, results for functions depending on several parameters are expected to be much less encouraging than in the 1-parameter case. This is indeed true. For example, analytic symmetric functions in 2-parameters do not even admit differentiable eigenvalues, as the next example illustrates.

Example 1.4.3. [18] Let $(x_1, x_2) \in \mathbb{R}^2$, and consider the function

$$A(x_1, x_2) = \begin{bmatrix} x_1 & x_2 \\ x_2 & -x_1 \end{bmatrix} .$$

The eigenvalues are $\pm\sqrt{x_1^2 + x_2^2}$, which are not differentiable at $(0, 0)$ (exactly the eigenvalues pictured in Figure 1). Of course, the problem is the lack of global differentiability at the origin, where both eigenvalues are 0. We notice that viewing $A(x_1, x_2)$ as a function of one parameter (holding the other frozen), renders analytic eigenvalues.

Example 1.4.3 notwithstanding, an important and useful tool in our investigation of smooth matrix valued functions in two-parameters is the "block-diagonalization" result of Hsieh and Sibuya, and Gingold, [17] and [11]. This result allows to focus locally, in the neighborhood of coalescing eigenvalues or singular values.

Theorem 1.4.4 (Block-Diagonalization)**.** *Let R be a closed rectangular region in $\mathbb{R}^2$. Suppose that the eigenvalues of $A \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$, $e \geq 0$, can be labeled so that they belong to two disjoint sets for all $x \in R$: $\lambda_1(x), \ldots, \lambda_p(x)$ in $\Lambda_1(x)$ and $\lambda_{p+1}(x), \ldots, \lambda_n(x)$ in $\Lambda_2(x)$, $\Lambda_1(x) \cap \Lambda_2(x) = \emptyset$, $\forall x \in R$. Further, assume that complex conjugate eigenvalues are put in the same group. Then, there exists $M \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$, invertible, such that*

$$M^{-1}(x)A(x)M(x) =: S = \begin{bmatrix} S_1(x) & 0 \\ 0 & S_2(x) \end{bmatrix} , \forall x \in R ,$$

where $S_1 \in \mathcal{C}^e(R, \mathbb{R}^{p\times p})$, $S_2 \in \mathcal{C}^e(R, \mathbb{R}^{(n-p)\times(n-p)})$, and the eigenvalues of $S_i(x)$ are those in $\Lambda_i(x)$, for all $x \in R$ and $i = 1, 2$.

We notice that the function M is not unique, in general.

A useful consequence of Theorem 1.4.4 is the following result, which we state as a Corollary.

Corollary 1.4.5. *Let $M \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$ be the function of which in Theorem 1.4.4. Let Γ be a simple closed curve in R, parametrized as a $\mathcal{C}^p$ ($p \geq 0$) function γ in the variable t, so that the function $\gamma : \ t \in \mathbb{R} \to R$ is $\mathcal{C}^p$ and 1-periodic. Let $m = \min(e, p)$, and let M_γ be the $\mathcal{C}^m$ function $M(\gamma(t))$, $t \in \mathbb{R}$. Then, we have $M_\gamma \in \mathcal{C}^m_1(\mathbb{R}, \mathbb{R}^{n\times n})$.*

Proof. The result is immediate upon considering the composite function M_γ and using the stated smoothness and periodicity results. □

Remarks 1.4.6.

(i) Naturally, Theorem 1.4.4 can be refined to any number of disjoint groups of eigenvalues. In the limiting case, if all eigenvalues are distinct, Theorem 1.4.4 says that we can find a $\mathcal{C}^e$ basis of real eigenvectors, and S is diagonal, with (2×2) bumps along the diagonal corresponding to complex conjugate eigenvalues.

(ii) In case A is also symmetric, which will be the case of interest for us, then M can be taken orthogonal and S stays symmetric. In this case, if the eigenvalues are distinct in R, then the orthogonal function M has diagonalized A: One has a $\mathcal{C}^e$ Schur decomposition. Naturally, in this case, M is essentially unique: the degree of non-uniqueness is solely determined by the ordering of the eigenvalues and the signs of the columns of M. Further, Corollary 1.4.5 will give M_γ as a 1-periodic function.

The following result is another useful consequence of Theorem 1.4.4, and clarifies the degree of non-uniqueness in the block-diagonalization result for symmetric functions. We state it as a Theorem (see [6, Lemma 2.2 and Theorem 2.3]).

Theorem 1.4.7. *Let $A \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$, $e \geq 0$, be symmetric. Suppose that the eigenvalues of A can be labeled so that they belong to two disjoint sets for all $x \in R$: $\lambda_1(x), \ldots, \lambda_p(x)$ in $\Lambda_1(x)$ and $\lambda_{p+1}(x), \ldots, \lambda_n(x)$ in $\Lambda_2(x)$, $\Lambda_1(x) \cap \Lambda_2(x) = \emptyset$, $\forall x \in R$.*

Consider $Q \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$, orthogonal, guaranteed to exist by Theorem 1.4.4, such

that

$$Q^T(x)A(x)Q(x) =: S = \begin{bmatrix} S_1(x) & 0 \\ 0 & S_2(x) \end{bmatrix} , \forall x \in R ,$$

where $S_1(x) = S_1^T(x) \in \mathbb{R}^{p\times p}$, $S_2(x) = S_2^T(x) \in \mathbb{R}^{(n-p)\times(n-p)}$, *for all* $x \in R$, *and the eigenvalues of* $S_i(x)$ *are those in* $\Lambda_i(x)$, *for all* $x \in R$ *and* $i = 1, 2$.

Then, any other $\mathcal{C}^e$ *orthogonal function* U *achieving a block diagonalization of* A *in two groups corresponding to the eigenvalues in* Λ_1, Λ_2, *must have the form:*

$$U(x) \;=\; Q(x) \begin{bmatrix} V_1(x) & 0 \\ 0 & V_2(x) \end{bmatrix} ,$$

where the $\mathcal{C}^e$ *functions* V_1 *and* V_2 *are orthogonal, taking values in* $\mathbb{R}^{p\times p}$ *and* $\mathbb{R}^{(n-p)\times(n-p)}$ *respectively.*

Proof. The proof is a consequence of the fact that orthonormal bases of invariant subspaces associated to disjoint groups of eigenvalues of symmetric functions are mutually orthogonal. In fact, consider X_1, X_2, with orthonormal columns and such that

$$AX_1 = X_1S_1, \;\; AX_2 = X_2S_2 \,. \tag{1}$$

We need to show that $X_2^TX_1 = 0$. Multiplying equations (1), respectively, by X_2^T and X_1^T, and using the symmetry of A and S_i, $i = 1, 2$, we have:

$$X_2^TX_1S_1 = X_2^TAX_1 = \big(X_1^TAX_2\big)^T = \big(X_1^TX_2S_2\big)^T = S_2X_2^TX_1 \,.$$

Using the last equation, we derive the following Lyapunov equation for $X_2^TX_1$:

$$\big(X_2^TX_1\big)S_1 - S_2\big(X_2^TX_1\big) = 0 \,,$$

which admits the unique solution $X_2^TX_1 = 0$ since the spectra of S_1 and S_2 are disjoint[2].

[2]In general, the matrix equation $AX + XB = C$ has a unique solution for any C if and only if $\lambda_i(A) + \lambda_j(B) \neq 0$ for all i, j

Now, if U is an orthogonal function having achieved the block reduction of which in the theorem, writing $Q = [Q_1, Q_2]$ and $U = [U_1, U_2]$ with the partitioning inherited by the dimensions of the eigenvalues' groups, we must have that U_i and Q_i, $i = 1, 2$, span the same subspace and are thus related as stated. □

Again, Theorem 1.4.7 can be refined to any number of disjoint blocks of eigenvalues as well.

To properly characterize periodicity of decompositions, we will also need the Lemma below, which is a generalization of an example given by Sibuya in [31]. Its main point is that, for a given continuous 1-periodic matrix function A, continuous 1-periodic and 2-periodic decompositions cannot coexist. We stress that this is true since we take the factors to be real-valued. In a somewhat more general fashion than the previous cases, we give the result for a general eigendecomposition. (A similar result holds for the SVD as well; the extension is straightforward and, therefore, omitted.)

Lemma 1.4.8. *Let $A \in \mathcal{C}_1^0(\mathbb{R}, \mathbb{R}^{n\times n})$ be such that*

$$A(t) = S(t)\Lambda(t)S^{-1}(t) \ , \ \forall t \ ,$$

with:

(i) $\Lambda \in \mathcal{C}_1^0(\mathbb{R}, \mathbb{R}^{n\times n})$ diagonal with distinct diagonal entries,

(ii) $S \in \mathcal{C}_2^0(\mathbb{R}, \mathbb{R}^{n\times n})$ invertible, with

$$S(t+1) = S(t)\ D \ , \ \forall t \in \mathbb{R} \ ,$$

where D is diagonal with $D_{ii} = \pm 1$ for all i, but $D \neq I_n$.

Then, there is no matrix function T such that:

$T \in \mathcal{C}_1^0(\mathbb{R}, \mathbb{R}^{n\times n})$, invertible, and $T(t)A(t)T^{-1}(t) = \Lambda(t)$ for all $t \in \mathbb{R}$.

Proof. Assume that such a function T exist. Then we have

$$\Lambda(t)T(t)S(t) = T(t)S(t)\Lambda(t)\,,$$

for all $t \in \mathbb{R}$. Since $\Lambda(t)$ has distinct diagonal entries for all $t \in \mathbb{R}$, it follows that $B(t) := T(t)S(t)$ is diagonal for all $t \in \mathbb{R}$. Let us denote its diagonal entries by $b_1(t), \ldots, b_n(t)$, for all t. Being $B(t)$ nonsingular for all $t \in \mathbb{R}$, the scalar functions b_i never vanish. On the other hand, we have $B(t+1) = T(t+1)S(t+1) = T(t)S(t)D = B(t)D$, for all $t \in \mathbb{R}$, hence there must exist an index i for which $b_i(t+1) = -b_i(t)$ for all $t \in \mathbb{R}$. This is a contradiction, because the functions b_i are continuous for $t \in \mathbb{R}$. □

The result below concerns the smoothness of the polar decomposition of an invertible matrix, and will be used in Section 2.2.1.2. Although the proof is simple, the result is of interest.

Lemma 1.4.9. *Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$. Assume that $A_0 := A(\xi_0)$ is invertible. Then, there is an open neighborhood of ξ_0, call it Ω_0, where A admits the polar factorization $A = QP$, where $Q \in \mathcal{C}^e(\Omega_0, \mathbb{R}^{n\times n})$ is orthogonal, and $P \in \mathcal{C}^e(\Omega_0, \mathbb{R}^{n\times n})$ is symmetric positive definite.*

Proof. It is well known that the matrix A_0, being invertible, admits a unique polar factorization $A_0 = Q_0P_0$, with Q_0 orthogonal and P_0 symmetric positive definite; e.g., see [13]. The matrix P_0 is the unique positive definite square root of $A_0^TA_0$. Since the function A is invertible in a neighborhood Ω_0 of ξ_0, then it admits a unique polar factorization at each given $\xi \in \Omega_0$. That the factors are smooth in a neighborhood of ξ_0 is a consequence of the implicit function theorem. To witness, consider the matrix equation for the square root P: $F(P) := P^2 - A^TA = 0$. The Frechét derivative of F at P_0 is the linear operator $F'(P_0) :\ X \to P_0X + XP_0$, which is invertible, since P_0 is positive definite (see footnote at page 14). Therefore, there is a unique $\mathcal{C}^e$ (symmetric

positive definite) square root P of A^TA in a neighborhood of ξ_0, and passing through P_0, and hence a unique $\mathcal{C}^e$ polar factorization, by letting $Q = AP^{-1}$. □

Interesting results that concern eigenvectors of parameter-dependent matrices have appeared in the Chemistry and Physics literature, particularly in the case of matrices with a *degenerate* spectrum (i.e., eigenvalues of multiplicity higher than 1). In 1963, theoretical chemists Herzberg and Longuet-Higgins ([15]) reported that the (real) eigenfunctions of a 2×2 real symmetric matrix function change sign under smooth continuation around a conical intersection. In 1984, the physicist Berry ([3]) introduced the so-called *geometric phase*, a phase $e^{i\gamma_n(C)}$ acquired by the n-th eigenstate of a quantal system as the parameter-dependent (generally complex) Hamiltonian $\widehat{H}(R)$ of the system is slowly varied around a closed circuit C in parameter space. If C lies closed to a degeneracy of the spectrum and the Hamiltonian $\widehat{H}$ is real symmetric, $\gamma_n(C)$ takes a simple form which contains, as a special case, the sign-change of the eigenfunctions discovered by Herzberg and Longuet-Higgins. In both the above-cited works, the mathematical arguments used appear to be local and case specific. Two results that will be given in Chapter 2 (i.e., Theorems 2.1.2 and 2.1.8) are similar in spirit to those reported in the above cited works, but are given in their most general form, as well as with a solid mathematical proof.

The major merits of this thesis are:

(i) mathematical formalization and rigorous proof of some known results concerning real symmetric matrix functions with one pair of coalescing eigenvalues,

(ii) extension to the case of several coalescing points (for the same pair or for distinct pairs of eigenvalues) and definition of coalescing of multiplicity higher than 1,

(iii) study of the SVD case,

(iv) derivation of theoretical results that allow the development of numerical techniques global in nature (whereas the numerical approach suggested in [33] appears to be local),

(v) development and implementation of algorithms to locate coalescing points that always work with the full $(n \times n)$ problem and do not pose any restriction on the geometry of the domain.

1.5 Background in Differential Geometry

In the remainder of this introduction we give some results from differential geometry which are needed to justify our later assumptions. We refer to [16] for background on these concepts. First of all, we recall the *Regular Value Theorem* (see [16, Theorem 3.2, p.22]) in a simplified form sufficient for our purposes.

Theorem 1.5.1. [Regular Value Theorem] *If $f : \Omega \to \mathbb{R}$ is a $\mathcal{C}^e$ map, $e \geq 1$, and 0 is a regular value of f, then $f^{-1}(0)$ is a $\mathcal{C}^e$ sub-manifold of Ω. In other words, the set $\{x \in \Omega \ : \ f(x) = 0\}$ is a (collection of) $\mathcal{C}^e$ curve(s) in Ω.*

Remark 1.5.2. The assumption that 0 be a regular value of f translates into the requirement that at values $x \in \Omega$, where $f(x) = 0$, we have $\nabla f(x) \neq 0$. We notice that (in case $e \geq 2$) the Morse-Sard theorem (see [16, Theorem 1.3, p.69]) tells us that we should expect that 0 be a regular value. Also, notice that Theorem 1.5.1 allows for $f^{-1}(0)$ to be given by the union of several non-intersecting curves, which may be either closed or extend forever; of course, given ξ_0 such that $f(\xi_0) = 0$, there is a unique curve through ξ_0.

We will also need the concept of *transversal intersection* of two smooth curves.

Definition 1.5.3. When two $\mathcal{C}^e$, $e \geq 1$, curves in $\Omega \subseteq \mathbb{R}^2$ intersect each other at a point ξ_0, we call the intersection transversal if the tangent vectors to the two curves at ξ_0 are not multiple of each other.

In the cases of interest to us, transversal intersection can be characterized as follows.

We have two $\mathcal{C}^e$, $e \geq 1$, functions from Ω to $\mathbb{R}$, f_1 and f_2, which vanish at a point $\xi_0 \in \Omega$: $f_1(\xi_0) = f_2(\xi_0) = 0$, and $\nabla f_1(\xi_0) \neq 0$, $\nabla f_2(\xi_0) \neq 0$. Assume that 0 is a regular value for f_1 and f_2. Thus, by Theorem 1.5.1, there are two well defined smooth curves through ξ_0: $\{x \in \Omega \ : \ f_1(x) = 0\}$ and $\{x \in \Omega \ : \ f_2(x) = 0\}$. Transversal intersection of these two curves at ξ_0 means that the Jacobian $\begin{bmatrix} \nabla f_1(x) \\ \nabla f_2(x) \end{bmatrix}$ is invertible at ξ_0: $\begin{bmatrix} \nabla f_1(\xi_0) \\ \nabla f_2(\xi_0) \end{bmatrix}$ invertible. As it is well known, transversal intersection of two smooth curves is a *generic* property. Thus, if two curves intersect non-transversally, there is an arbitrarily small perturbation for which the two curves will intersect transversally (or not at all).

In our results in Sections 2.1 and 2.2, we will make assumptions which reflect precisely generic properties of $\mathcal{C}^e$ matrix valued functions A.

CHAPTER II

THEORETICAL RESULTS: TWO PARAMETERS

In this chapter we state and prove several theoretical results that concern matrix valued functions that depend on two parameters. Our main theoretical result in this thesis is condensed in the final theorem we give at the end of the chapter, Theorem 2.2.13. This theorem is analogous to the Intermediate Value Theorem in Calculus, where one is able to infer that a continuous function has a zero in a closed interval whenever the function has opposite signs at the endpoints of this interval. Here, we will have a two-dimensional analog of this result, retaining the original topological flavor: Oversimplifying it, Theorem 2.2.13 says that there is a coalescing inside a curve if a certain continuous function changes sign along the curve. The function which one needs to monitor is obtained by the continuous orthogonal factors in the decomposition of the two-parameter matrix function restricted to the curve. We will prove Theorem 2.2.13 proceeding by examining simpler cases first, and then showing that more complicated scenarios can be brought back to these simpler cases. Finally, we remark here that our theoretical result will lend very nicely to a numerical approach to locate coalescing points, as we have reported in Chapter 4.

2.1 One Pair Coalescing

We first tackle the case of coalescing eigenvalues of symmetric matrix functions. The situation of coalescing singular values of a general function will follow from this case.

2.1.1 Coalescing Eigenvalues of Symmetric Matrices

Our first result is when $A(x)$ takes values in $\mathbb{R}^{2\times 2}$. This "simpler" case will turn out to be the stepping stone for the general case. Moreover, this (2×2) case already

presents the key essential features, so we will be able to present the fundamental ideas in a transparent setting. First, we have this elementary result.

Lemma 2.1.1. *Let* $P = \begin{bmatrix} a & b \\ b & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$ *be a given symmetric matrix. Then, this matrix has two identical eigenvalues if and only if* $P = \lambda I$.

Proof. Since the eigenvalues are given by $\frac{1}{2}\big(a + d \pm \sqrt{(a-d)^2 + 4b^2}\big)$, they coincide if and only if $a - d = 0$ and $b = 0$. □

We now have

Theorem 2.1.2 (Symmetric (2×2) case)**.** *Consider* $P \in \mathcal{C}^e(\Omega, \mathbb{R}^{2\times 2})$, $e \geq 1$, *symmetric. For all* $x \in \Omega$, *write*

$$P(x) = \begin{bmatrix} a(x) & b(x) \\ b(x) & d(x) \end{bmatrix},$$

and let λ_1 *and* λ_2 *be its two continuous eigenvalues, labeled so that* $\lambda_1(x) \geq \lambda_2(x)$ *for all* x *in* Ω. *Assume that there exists a unique point* $\xi_0 \in \Omega$ *where the eigenvalues coincide:* $\lambda_1(\xi_0) = \lambda_2(\xi_0)$. *Consider the* $\mathcal{C}^e$ *function* $F : \Omega \to \mathbb{R}^2$ *given by*

$$F(x) = \begin{bmatrix} a(x) - d(x) \\ b(x) \end{bmatrix}, \tag{2}$$

and assume that 0 *is a regular value for both functions* $a - d$ *and* b. *Then, consider the two* $\mathcal{C}^e$ *curves* Γ_1 *and* Γ_2 *through* ξ_0, *given by the zero-set of the components of* F: $\Gamma_1 = \{x \in \Omega : a(x) - d(x) = 0\}$, $\Gamma_2 = \{x \in \Omega : b(x) = 0\}$. *Assume that* Γ_1 *and* Γ_2 *intersect transversally at* ξ_0.

Let Γ *be a simple closed curve*[1] *enclosing the point* ξ_0, *and let it be parametrized as a* $\mathcal{C}^p$ $(p \geq 0)$ *function* γ *in the variable* t, *so that the function* $\gamma : t \in \mathbb{R} \to \Omega$ *is*

[1] *Also called a Jordan curve*

$\mathcal{C}^p$ *and* 1*-periodic. Let* $m = \min(e, p)$*, and let* P_γ *be the* $\mathcal{C}^m$ *function* $P(\gamma(t))$*,* $t \in \mathbb{R}$*. Then, for all* $t \in \mathbb{R}$*,* $P_\gamma(t)$ *has the eigendecomposition*

$$P_\gamma(t) = V_\gamma(t)\Lambda_\gamma(t)V_\gamma^T(t)$$

such that:

(i) $\Lambda_\gamma \in \mathcal{C}_1^m(\mathbb{R}, \mathbb{R}^{2\times 2})$ *and diagonal:* $\Lambda_\gamma(t) = \begin{bmatrix} \lambda_1(\gamma(t)) & 0 \\ 0 & \lambda_2(\gamma(t)) \end{bmatrix}$ *for all* $t \in \mathbb{R}$*;*

(ii) $V_\gamma \in \mathcal{C}_2^m(\mathbb{R}, \mathbb{R}^{2\times 2})$ *real orthogonal, and* $V_\gamma(t+1) = -V_\gamma(t)$ *for all* $t \in \mathbb{R}$*.*

Proof. We remark that, because of Lemma 2.1.1,

$$\lambda_1(x) = \lambda_2(x) \iff F(x) = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

and, by hypothesis, ξ_0 is the unique root of $F(x)$ in Ω.

The proof will go as follows. First, we will prove that the asserted results hold true along a small circle C around ξ_0. Then, we will show that the same periodicity results hold when we continuously deform C into Γ.

Let C be a circle centered at ξ_0, of radius small enough so that the circle goes through each of Γ_1 and Γ_2 at exactly two distinct points (see Figure 3). This is possible since Γ_1 and Γ_2 intersect transversally at ξ_0. Further, let C be parametrized by a continuous 1-periodic function[2] ρ, $\rho(t+1) = \rho(t)$, for all $t \in \mathbb{R}$.

Now, let us consider $P_\rho(t) = P(\rho(t))$, $t \in \mathbb{R}$, which is continuous and 1-periodic, with distinct eigenvalues, so that its eigenvalues are necessarily 1-periodic. Also, the eigenvectors of $P_\rho(t)$, call them $V_\rho(t)$, are uniquely determined (for each t) up to sign. The first column of $V_\rho(\cdot)$ is given by an orthonormal basis for $\mathrm{Ker}(P_\rho(\cdot) - \lambda_1(\rho(\cdot))I)$. The function $(P_\rho(\cdot) - \lambda_1(\rho(\cdot))I)$ is a continuous 1-periodic constant rank matrix,

[2]this parametrization does not need to be specified yet; see Remark 2.1.5

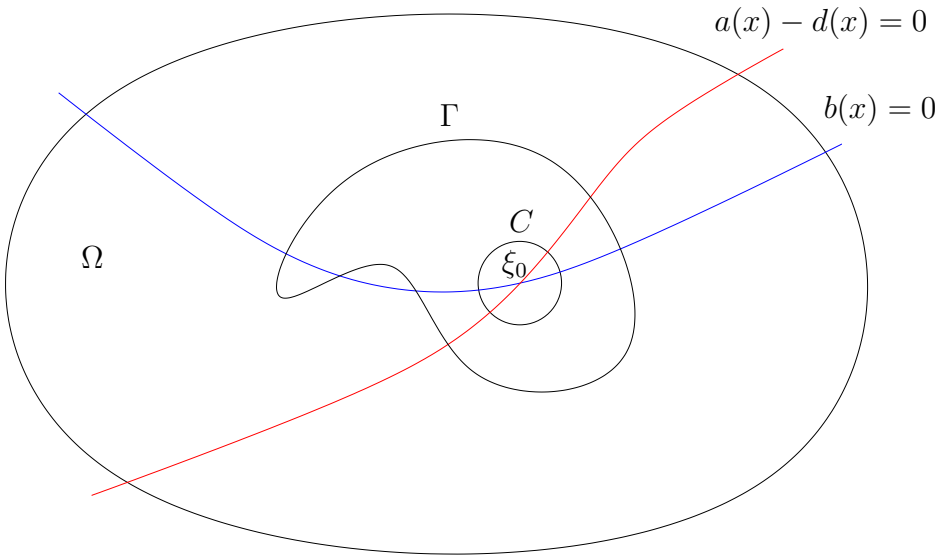

Figure 3: Transversal Intersection at ξ_0

for which the existence of a continuous periodic Schur decomposition, possibly with 2-periodic eigenvectors, was proved in [6]. We can therefore write:

$$\left[P_\rho(t) - \lambda_1(\rho(t))I\right] \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} = (\lambda_2(\rho(t)) - \lambda_1(\rho(t))) \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \ , \ t \in \mathbb{R} \qquad (3)$$

where the continuous scalar valued functions u_1, u_2 have period either 1 or 2. Let us use the notation $a_\rho(\cdot) = a(\rho(\cdot))$, and same for $b(\cdot)$ and $d(\cdot)$. Then, after a simple computation, equation (3) yields:

$$\begin{cases} \left(\frac{a_\rho(t)-d_\rho(t)}{2} + \frac{1}{2}\sqrt{(a_\rho(t)-d_\rho(t))^2 + 4b_\rho(t)^2}\,\right) u_1(t) = -b_\rho(t)\, u_2(t) \\ \left(\frac{d_\rho(t)-a_\rho(t)}{2} + \frac{1}{2}\sqrt{(a_\rho(t)-d_\rho(t))^2 + 4b_\rho(t)^2}\,\right) u_2(t) = -b_\rho(t)\, u_1(t) \end{cases} , \ t \in \mathbb{R} \ .$$

From these last equations, it follows that u_1 (respectively, u_2) changes sign if and only if b_ρ goes through zero and $(a_\rho - d_\rho) > 0$ (respectively, $(a_\rho - d_\rho) < 0$). Therefore, each of the two functions u_1 and u_2 changes sign only once over any interval of length 1. But no continuous function of period 1 can change sign only once over one period. Therefore, u_1 and u_2 must be 2-periodic functions and the periodicity assertions of

the theorem follow relatively to the curve $\rho(t)$: $V_\rho \in \mathcal{C}_2(\mathbb{R}, \mathbb{R}^{2\times 2})$, where

$$V_\rho(t) = \begin{bmatrix} -u_2(t) & u_1(t) \\ u_1(t) & u_2(t) \end{bmatrix} \ , \ \text{ or } \ V_\rho(t) = \begin{bmatrix} u_2(t) & u_1(t) \\ -u_1(t) & u_2(t) \end{bmatrix} \ , \ t \in \mathbb{R} \ .$$

Now, consider the case of Γ. As before, we notice that the existence of a $\mathcal{C}^m$ eigendecomposition for P along Γ, possibly with 2-periodic orthogonal factors, is known (e.g., see [6]). What we need to show is that the periodicity results that hold along the circle C also hold along the curve Γ.

Let us consider a homotopy $h(s,t)$, $(s,t) \in [0,1] \times [0,1]$ satisfying the following properties[3]:

(1) $h(s,t)$ is continuous in $(s,t) \in [0,1] \times [0,1]$ and for all $t \in [0,1]$: $h(0,t) = \rho(t)$, $h(1,t) = \gamma(t)$, and for any $s \in [0,1]$: $h(s,0) = h(s,1)$;

(2) h continuously (in s) deforms $\rho(\cdot)$ into $\gamma(\cdot)$ in such a way that –for each given $s \in (0,1)$– the simple closed curves $h(s,\cdot)$ are always contained in the interior of Γ and in the exterior of C.

Let us consider the function $P(h(s,t))$, $(s,t) \in [0,1] \times [0,1]$. $P(h(s,t))$ is continuous with distinct eigenvalues for all $(s,t) \in [0,1] \times [0,1]$. Therefore, by Theorem 1.4.4, we can write $P(h(s,t)) = V(s,t)\Lambda(s,t)V^T(s,t)$, where $\Lambda(s,t)$ and $V(s,t)$ are continuous, $\Lambda(s,t)$ is diagonal, and $V(s,t) = \begin{bmatrix} v_1(s,t) & v_2(s,t) \end{bmatrix}$ is real orthogonal. Let $f_k(s) = v_k^T(s,0)\, v_k(s,1)$, for $k = 1, 2$. Since $h(s,0) = h(s,1)$ for all $s \in [0,1]$, we have that f_1 and f_2 take values in $\{-1, 1\}$. Being continuous, they have to be constant over $[0,1]$. Therefore, we must have $f_1(0) = f_1(1) = -1$ and $f_2(0) = f_2(1) = -1$. This means that $V_\gamma(1) = -V_\gamma(0)$ and thus we obtain the asserted periodicity properties for V_γ. □

Remark 2.1.3. The eigendecomposition of P_γ in Theorem 2.1.2 is essentially unique, within the class of $\mathcal{C}^m$ Schur decompositions. The degree of non-uniqueness is given

[3]See remark 2.1.5

by the ordering of the eigenvalues and by the signs of the columns of V_γ. In particular, for $\mathcal{C}^m$–decompositions, the statement about periodicity holds unchanged.

Remark 2.1.4. We stress that the assumption of transversality for the curves Γ_1 and Γ_2 at ξ_0 is generic. With abuse of notation, we say that ξ_0 is a *generic coalescing point of eigenvalues* when Γ_1 and Γ_2 intersect transversally at ξ_0.

Remark 2.1.5. The existence of the homotopy $h(s,t)$ is non-trivial[4]. Let us call G the open, simply connected, region enclosed by Γ, and let us call D the closed disk with boundary C. Then, using [1, Theorem 14.25], there exists a homeomorphism $f : \overline{G - D} \to A$, where A is the annulus $A = \{x \in \mathbb{R}^2 \ : \ 1 \leq \|x\| \leq 2\}$, so that f maps C and Γ into the circles of radius 1 and 2 in A, respectively, call them C_1 and C_2. Now, we let $f(\gamma(t))$ be a parametrization of C_2. We can think of this parametrization as having the form $2\begin{bmatrix}\cos(\alpha(t))\\ \sin(\alpha(t))\end{bmatrix}$ with α continuous and such that $\alpha(0) = 0$ and $\alpha(t+1) = \alpha(t) + 2\pi$, for all $t \in \mathbb{R}$. We also parametrize C_1 as $\begin{bmatrix}\cos(\alpha(t))\\ \sin(\alpha(t))\end{bmatrix}$. Finally, we consider the pullback via f^{-1} of the linear homotopy in the annulus (i.e., $(1-s)C_1 + sC_2$), which will give the sought homotopy $h(s,t)$: $h(s,t) = f^{-1}\big((1+s)\begin{bmatrix}\cos(\alpha(t))\\ \sin(\alpha(t))\end{bmatrix}\big)$. [Notice that we can now qualify what parametrization we should choose for the circle C in the proof of Theorem 2.1.2: It is $\rho(t) = f^{-1}\big(\begin{bmatrix}\cos(\alpha(t))\\ \sin(\alpha(t))\end{bmatrix}\big)$.]

With the aid of Theorem 2.1.2, and of the reduction to block-diagonal form, we can now tackle the case of symmetric functions in $\mathbb{R}^{n\times n}$, whose eigenvalues coalesce at a unique point ξ_0. We will need the following definition (see Remark 2.1.4).

Definition 2.1.6. Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$ be a symmetric function with continuous eigenvalues $\lambda_1(x), \ldots, \lambda_n(x)$, $x \in \Omega$, satisfying

$$\lambda_1(x) > \lambda_2(x) > \ldots > \lambda_k(x) \geq \lambda_{k+1}(x) > \ldots > \lambda_n(x) \ , \ \forall x \in \Omega \ ,$$

and

$$\lambda_k(x) = \lambda_{k+1}(x) \Longleftrightarrow x = \xi_0 \in \Omega \ .$$

[4]We thank Prof. M. Ghomi of Georgia Tech for clarifying to us the existence of this homotopy

Let R be a rectangular region $R \subseteq \Omega$ containing ξ_0 in its interior. Moreover, let

(1) $Q \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$ be a $\mathcal{C}^e$ orthogonal function achieving the reduction of which in Theorem 1.4.4 (and see Remarks 1.4.6)

$$Q^T(x)A(x)Q(x) = \begin{bmatrix} \Lambda_1(x) & 0 & 0 \\ 0 & P(x) & 0 \\ 0 & 0 & \Lambda_2(x) \end{bmatrix}, \forall x \in R ,$$

where $\Lambda_1 \in \mathcal{C}^e(R, \mathbb{R}^{(k-1)\times(k-1)})$ and $\Lambda_2 \in \mathcal{C}^e(R, \mathbb{R}^{(n-k-1)\times(n-k-1)})$, such that, for all $x \in R$, $\Lambda_1(x) = \mathrm{diag}(\lambda_1(x), \ldots, \lambda_{k-1}(x))$, and $\Lambda_2(x) = \mathrm{diag}(\lambda_{k+2}(x), \ldots, \lambda_n(x))$. Moreover, $P \in \mathcal{C}^e(R, \mathbb{R}^{2\times 2})$ is symmetric, and $P(x)$ has eigenvalues $\lambda_k(x), \lambda_{k+1}(x)$ for each $x \in R$;

(2) for all $x \in R$, write $P(x) = \begin{bmatrix} a(x) & b(x) \\ b(x) & d(x) \end{bmatrix}$, and define the function F and the curves Γ_1 and Γ_2 as in Theorem 2.1.2, for $x \in R$.

Then, we call ξ_0 a *generic coalescing point of eigenvalues* in Ω, if the curves Γ_1 and Γ_2 intersect transversally at ξ_0.

Before proceeding, we must justify Definition 2.1.6, since the $\mathcal{C}^e$ function Q which does the block diagonalization of which in Theorem 1.4.4 is not unique. In particular, the function P in Definition 2.1.6 is not unique, and we need to argue that any other possible function would have shared the same property. This is true, and it is the content of the following result.

Theorem 2.1.7. *Let A be as in Definition 2.1.6 and let Q be a fixed $\mathcal{C}^e$ orthogonal function achieving the reduction of which in Theorem 1.4.4:*

$$Q^T(x)A(x)Q(x) = \begin{bmatrix} \Lambda_1(x) & 0 & 0 \\ 0 & P(x) & 0 \\ 0 & 0 & \Lambda_2(x) \end{bmatrix}, \forall x \in R ,$$

as in Definition 2.1.6. Write $P(x) = \begin{bmatrix} a(x) & b(x) \\ b(x) & d(x) \end{bmatrix}$. *Let* $U \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$ *be another orthogonal function achieving a similar block reduction:*

$$U^T(x)A(x)U(x) = \begin{bmatrix} \Lambda_1(x) & 0 & 0 \\ 0 & \widetilde{P}(x) & 0 \\ 0 & 0 & \Lambda_2(x) \end{bmatrix} , \forall x \in R .$$

Write $\widetilde{P}(x) = \begin{bmatrix} \widetilde{a}(x) & \widetilde{b}(x) \\ \widetilde{b}(x) & \widetilde{d}(x) \end{bmatrix}$.

Then, if the curves $\Gamma_1 = \{x \in R : (a-d)(x) = 0\}$ *and* $\Gamma_2 = \{x \in R : b(x) = 0\}$ *intersect transversally at* ξ_0*, so do the curves* $\widetilde{\Gamma}_1 = \{x \in R : (\widetilde{a} - \widetilde{d})(x) = 0\}$ *and* $\widetilde{\Gamma}_2 = \{x \in R : \widetilde{b}(x) = 0\}$.

Proof. Obviously, P and $\widetilde{P}$ are similar, and thus have same eigenvalues, which coalesce only at ξ_0, so that $\widetilde{\Gamma}_1$ and $\widetilde{\Gamma}_2$ intersect at ξ_0, and only there. By using Theorem 1.4.7, we know that $\widetilde{P}(x) = V^T(x)P(x)V(x)$, where $V \in \mathcal{C}^e(R, \mathbb{R}^{2\times 2})$ is orthogonal.

Write $V(x) = [v_1(x), v_2(x)]$, for all $x \in R$. From the fact that $v_i^T(x)v_i(x) = 1$, $\forall x \in R$, it follows that

$$\text{(a)} \qquad v_i^T(x)\big(\partial_{x_j} v_i(x)\big) \;=\; 0 \;,\; i,j = 1,2,\; \forall x \in R\,.$$

Similarly, from the fact that $v_1^T(x)v_2(x) = 0$, $\forall x \in R$, it follows that

$$\text{(b)} \qquad v_1^T(x)\big(\partial_{x_j} v_2(x)\big) \;=\; -v_2^T(x)\big(\partial_{x_j} v_1(x)\big) \;,\; j = 1,2,\; \forall x \in R\,.$$

As a consequence of (a)-(b), and since $P(\xi_0)$ is a scalar multiple of the identity, direct computation shows that

$$\nabla\widetilde{b}(\xi_0) \;=\; \Big[v_1^T(\partial_{x_1}P)v_2 \qquad v_1^T(\partial_{x_2}P)v_2\Big]_{\xi_0} ,$$

and

$$\nabla(\widetilde{a} - \widetilde{d})(\xi_0) \;=\; \Big[v_1^T(\partial_{x_1}P)v_1 - v_2^T(\partial_{x_1}P)v_2 \qquad v_1^T(\partial_{x_2}P)v_1 - v_2^T(\partial_{x_2}P)v_2\Big]_{\xi_0} .$$

Now, since $V(\xi_0)$ is orthogonal, it must have one of the two forms (for some α):

$$\text{(i)} \quad V(\xi_0) = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{bmatrix} \qquad \text{or} \qquad \text{(ii)} \quad V(\xi_0) = \begin{bmatrix} \cos(\alpha) & \sin(\alpha) \\ \sin(\alpha) & -\cos(\alpha) \end{bmatrix} .$$

In case (i), the Jacobian matrix $\begin{bmatrix} \nabla\widetilde{b} \\ \nabla(\widetilde{a}-\widetilde{d}) \end{bmatrix}_{\xi_0}$ rewrites as

$$\begin{bmatrix} \nabla\widetilde{b} \\ \nabla(\widetilde{a}-\widetilde{d}) \end{bmatrix}_{\xi_0} = \begin{bmatrix} \cos(2\alpha) & \sin(2\alpha) \\ -2\sin(2\alpha) & 2\cos(2\alpha) \end{bmatrix} \begin{bmatrix} \partial_{x_1} b & \partial_{x_2} b \\ \frac{1}{2}\partial_{x_1}(a-d) & \frac{1}{2}\partial_{x_2}(a-d) \end{bmatrix}_{\xi_0} , \quad (4)$$

which is nonsingular since the Jacobian $\begin{bmatrix} \partial_{x_1} b & \partial_{x_2} b \\ \partial_{x_1}(a-d) & \partial_{x_2}(a-d) \end{bmatrix}_{\xi_0}$ is nonsingular (Γ_1 and Γ_2 intersect transversally at ξ_0). In case (ii), the Jacobian matrix $\begin{bmatrix} \nabla\widetilde{b} \\ \nabla(\widetilde{a}-\widetilde{d}) \end{bmatrix}_{\xi_0}$ rewrites as

$$\begin{bmatrix} \nabla\widetilde{b} \\ \nabla(\widetilde{a}-\widetilde{d}) \end{bmatrix}_{\xi_0} = \begin{bmatrix} -\cos(2\alpha) & \sin(2\alpha) \\ 2\sin(2\alpha) & 2\cos(2\alpha) \end{bmatrix} \begin{bmatrix} \partial_{x_1} b & \partial_{x_2} b \\ \frac{1}{2}\partial_{x_1}(a-d) & \frac{1}{2}\partial_{x_2}(a-d) \end{bmatrix}_{\xi_0} , \quad (5)$$

which is likewise nonsingular. Therefore, $\widetilde{\Gamma}_1$ and $\widetilde{\Gamma}_2$ intersect transversally at ξ_0. □

Of course, for a coalescing point of eigenvalues to be a generic coalescing point is a generic property.

We can now give the following result for symmetric $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$.

Theorem 2.1.8. *Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$ be symmetric. Let $\lambda_1(x), \ldots, \lambda_n(x)$, $x \in \Omega$, be its continuous eigenvalues. Suppose that*

$$\lambda_1(x) > \lambda_2(x) > \ldots > \lambda_k(x) \geq \lambda_{k+1}(x) > \ldots > \lambda_n(x) \ , \forall x \in \Omega \ ,$$

and

$$\lambda_k(x) = \lambda_{k+1}(x) \Longleftrightarrow x = \xi_0 \in \Omega \ .$$

Let ξ_0 be a generic coalescing point.

Let Γ be a simple closed curve in Ω enclosing the point ξ_0, and let it be parametrized as a $\mathcal{C}^p$ ($p \geq 0$) function γ in the variable t, so that the function $\gamma : \ t \in \mathbb{R} \to \Omega$ is $\mathcal{C}^p$ and 1-periodic. Let $m = \min(e, p)$, and let A_γ be the $\mathcal{C}^m$ function $A(\gamma(t))$, $t \in \mathbb{R}$.

Then, for all $t \in \mathbb{R}$, $A(\gamma(t))$ has the eigendecomposition

$$A_\gamma(t) = U_\gamma(t)\Lambda_\gamma(t)U_\gamma^T(t)$$

satisfying the following conditions:

(i) $\Lambda_\gamma \in \mathcal{C}_1^m(\mathbb{R}, \mathbb{R}^{n\times n})$ and diagonal: $\Lambda_\gamma(t) = \mathrm{diag}(\lambda_1(\gamma(t)), \ldots, \lambda_n(\gamma(t)))$, $\forall t \in \mathbb{R}$;

(ii) $U_\gamma \in \mathcal{C}_2^m(\mathbb{R}, \mathbb{R}^{n\times n})$ real orthogonal, and for all $t \in \mathbb{R}$

$$U_\gamma(t+1) = U_\gamma(t)D \ , \ \ D = \begin{bmatrix} I_{k-1} & 0 & 0 \\ 0 & -I_2 & 0 \\ 0 & 0 & I_{n-k-1} \end{bmatrix} .$$

Proof. Consider a rectangle $R \subseteq \Omega$ around ξ_0, and consider a function $Q \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$ giving the block decomposition

$$Q^T(x)A(x)Q(x) = \begin{bmatrix} \Lambda_1(x) & 0 & 0 \\ 0 & P(x) & 0 \\ 0 & 0 & \Lambda_2(x) \end{bmatrix} , \ \forall x \in R \ ,$$

as in Definition 2.1.6. Let C be a circle enclosing ξ_0 and contained in R, parametrized by a continuous 1-periodic function ρ, and let $P_\rho(t) = P(\rho(t))$, $t \in \mathbb{R}$; see Figure 4.

Let V_ρ be the orthogonal function of Theorem 2.1.2 associated to P_ρ, so that $V_\rho(t+1) = -V_\rho(t)$, for all t. Now, consider the following orthogonal continuous function

$$U_\rho(t) = Q(\rho(t)) \begin{bmatrix} I_{k-1} & 0 & 0 \\ 0 & V_\rho(t) & 0 \\ 0 & 0 & I_{n-k-1} \end{bmatrix} .$$

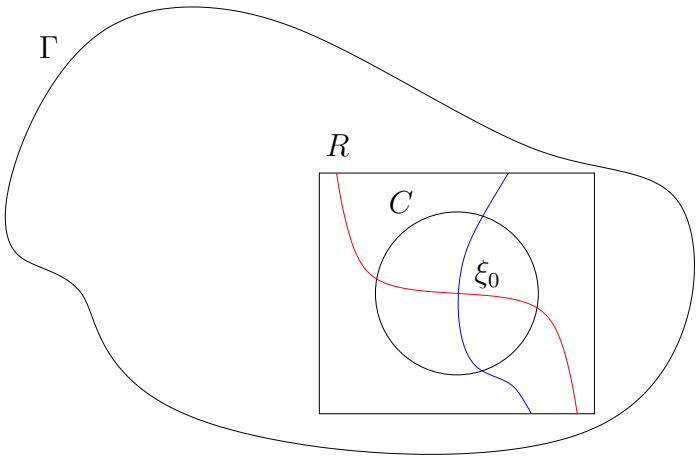

Figure 4: Generic coalescing at ξ_0.

Since $Q(\rho(t+1)) = Q(\rho(t))$ for all t (see Corollary 1.4.5), we then have

$$U_\rho(t+1) \;=\; U_\rho(t) \begin{bmatrix} I_{k-1} & 0 & 0 \\ 0 & -I_2 & 0 \\ 0 & 0 & I_{n-k-1} \end{bmatrix} .$$

We now need to argue that the same periodicity properties hold for U_γ. We do this in the same way as what we did in Theorem 2.1.2. Take a homotopy $h(s,t)$, $(s,t) \in [0,1]\times[0,1]$, as in Theorem 2.1.2, and consider the function $A(h(s,t))$, $(s,t) \in [0,1]\times[0,1]$. $A(h(s,t))$ is continuous with distinct eigenvalues for all $(s,t) \in [0,1]\times[0,1]$, and so –by Theorem 1.4.4– we can write $A(h(s,t)) = V(s,t)\Lambda(s,t)V^T(s,t)$, where $\Lambda(s,t)$ and $V(s,t)$ are continuous, $\Lambda(s,t)$ is diagonal, and $V(s,t)$ is real orthogonal. Partition V by columns: $V(s,t) = \begin{bmatrix} v_1(s,t) & \cdots & v_n(s,t) \end{bmatrix}$. Let $f_j(s) = v_j^T(s,0)\, v_j(s,1)$, for $j = 1,\ldots,n$. Since $h(s,0) = h(s,1)$ for all $s \in [0,1]$, we have that all f_j's take values in $\{-1,1\}$. Being continuous, we must have $f_j(0) = f_j(1) = 1$ for all $j \neq k, k+1$, and $f_j(0) = f_j(1) = -1$ for $j = k, k+1$, from which the result follows. □

Theorem 2.1.8 has an interesting geometric interpretation: the $\mathcal{C}^m$ eigenvectors of A_γ corresponding to the eigenvalues coalescing at ξ_0 (i.e., the k-th and $(k+1)$-st eigenvectors), "get upside down" as we complete one loop along the closed curve Γ. At the same time, it is worth stressing that the columns of U_γ (the eigenvectors)

associated to eigenvalues which do not coalesce in Ω maintain period 1, as A_γ has. Indeed, with the same technique used in the proof of Theorem 2.1.8, it is easy to refine Corollary 1.4.5 to obtain the result below, which shows that a continuous eigendecomposition along a simple curve Γ, not containing coalescing points (on or) inside it, has period 1.

Corollary 2.1.9. *Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$, and let let Γ be a simple closed curve in Ω, parametrized by the $\mathcal{C}^p$ and 1-periodic function γ. Let $m = \min(e, p)$, and let $A_\gamma \in \mathcal{C}^m_1(\mathbb{R}, \mathbb{R}^{n\times n})$ be the function $A(\gamma(t))$, $t \in \mathbb{R}$. If there are no coalescing points inside Γ (nor on it), then any $\mathcal{C}^m$ eigendecomposition of A_γ is 1-periodic.*

Remark 2.1.10. The block-diagonalization result in Theorem 1.4.4, and Lemma 2.1.1, are sufficient to describe, in essence, the idea of von Neumann and Wigner's proof of the non-crossing rule ([32]), see Section 1.2.1. Their proof is essentially based on the observation that, locally, in order to study the occurrence of coalescing of eigenvalues for a real symmetric $(n \times n)$ matrix A, it is enough to focus on a symmetric (2×2) matrix P, for which the condition of having two identical eigenvalues is equivalent to the following condition:

$$F(x) = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where P and F are defined as in Theorem 2.1.2. The equation above yields, in general, two independent constraints, hence they conclude that coalescing of eigenvalues is a codimension 2 phenomenon.

2.1.2 Coalescing Singular Values of General Matrices

Results similar to Theorem 2.1.2, Theorem 2.1.8, Theorem 2.1.7, and Corollary 2.1.9, hold true for the SVD of a matrix valued function A. To properly state these results, we need a preliminary result, similar to Theorem 1.4.4.

Theorem 2.1.11. *Let $R \subseteq \Omega$ be a rectangular region, and let $A \in \mathcal{C}^e(R, \mathbb{R}^{n\times n}), e \geq 0$. Suppose that the singular values of A can be labeled so that they belong to two disjoint sets for all $x \in \Omega$: $\sigma_1(x), \ldots, \sigma_p(x)$ in $\Sigma_1(x)$ and $\sigma_{p+1}(x), \ldots, \sigma_n(x)$ in $\Sigma_2(x)$, $\Sigma_1(x) \cap \Sigma_2(x) = \emptyset$, $\forall x \in R$. Then, there exist $U, V \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$, orthogonal, such that*

$$U^T(x)A(x)V(x) =: S = \begin{bmatrix} S_1(x) & 0 \\ 0 & S_2(x) \end{bmatrix} , \forall x \in R ,$$

where $S_1 \in \mathcal{C}^e(R, \mathbb{R}^{p\times p})$, $S_2 \in \mathcal{C}^e(R, \mathbb{R}^{(n-p)\times(n-p)})$, and the singular values of $S_i(x)$ are those in $\Sigma_i(x)$, for all $x \in R$ and $i = 1, 2$.

Proof. Using Theorem 1.4.4, for all $x \in R$ we have the two block-diagonalizations

$$U^T(x)A(x)A^T(x)U(x) = \begin{bmatrix} P_1(x) & 0 \\ 0 & P_2(x) \end{bmatrix} , \tag{6}$$

$$V^T(x)A^T(x)A(x)V(x) = \begin{bmatrix} R_1(x) & 0 \\ 0 & R_2(x) \end{bmatrix} , \tag{7}$$

where

(i) $U, V \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$ orthogonal,

(ii) $P_1, R_1 \in \mathcal{C}^e(R, \mathbb{R}^{p\times p})$ symmetric, and for all $x \in R$ the eigenvalues of $P_1(x)$ and $R_1(x)$ are the squares of the singular values in $\Sigma_1(x)$;

(iii) $P_2, R_2 \in \mathcal{C}^e(R, \mathbb{R}^{(n-p)\times(n-p)})$ symmetric, and for all $x \in R$ the eigenvalues of $P_2(x)$ and $R_2(x)$ are the squares of the singular values in $\Sigma_2(x)$.

For all $x \in R$, write

$$U^T(x)A(x)V(x) = \begin{bmatrix} S_1(x) & X_{12}(x) \\ X_{21}(x) & S_2(x) \end{bmatrix} ,$$

where the partitioning is inherited by the right hand sides of (6) and (7). Manipulation of (6) and (7) yields

$$U^T(x)A(x)A^T(x)A(x)V(x) =$$
$$= \begin{bmatrix} P_1(x) & 0 \\ 0 & P_2(x) \end{bmatrix} U^T(x)A(x)V(x) =$$
$$= U^T(x)A(x)V(x) \begin{bmatrix} R_1(x) & 0 \\ 0 & R_2(x) \end{bmatrix} ,$$

which implies that

$$P_1(x)X_{12}(x) = X_{12}(x)R_2(x)$$
$$P_2(x)X_{21}(x) = X_{21}(x)R_1(x) ,$$

for all x in R. By hypothesis, the spectra of P_1 and R_2, and of P_2 and R_1, are disjoint in R, and hence we must have $X_{12}(x) = 0$, $X_{21}(x) = 0$ (see footnote at page 14). □

Remarks 2.1.12.

(1) In general, just as in the case of Theorems 1.4.4 and 1.4.7, the block SVD of Theorem 2.1.11 is not unique. A result similar to Theorem 1.4.7 holds now as well, the proof being nearly identical to that of Theorem 1.4.7, and (with obvious notation) it can be phrased as follows.

Under the same assumptions of Theorem 2.1.11, suppose that U and $V \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$, are given, orthogonal functions (guaranteed to exist by Theorem 2.1.11), such that

$$U^T(x)A(x)V(x) =: S = \begin{bmatrix} S_1(x) & 0 \\ 0 & S_2(x) \end{bmatrix} , \forall x \in R ,$$

where the singular values of S_1 (respectively S_2) are those in Σ_1 (respectively Σ_2), for all $x \in R$. Then, any other $\mathcal{C}^e$ block SVD of A in R, in two groups corresponding to the singular values in Σ_1, Σ_2, must have the form:

$$U(x) \begin{bmatrix} W_1(x) & 0 \\ 0 & W_2(x) \end{bmatrix} , \quad V(x) \begin{bmatrix} Z_1(x) & 0 \\ 0 & Z_2(x) \end{bmatrix} ,$$

where the $\mathcal{C}^e$ functions W_1, Z_1, and W_2, Z_2, are orthogonal, taking values in $\mathbb{R}^{p\times p}$ and $\mathbb{R}^{(n-p)\times(n-p)}$ respectively.

(2) Of course, Theorem 2.1.11 can be refined to any number of groups of singular values which remain disjoint in R. In particular, if all singular values are distinct, S is diagonal. Moreover, if the singular values are distinct, then the $\mathcal{C}^e$ functions U and V are essentially unique: the degree of non-uniqueness is solely determined by the ordering of the diagonal entries of S and by joint changes of signs for the columns of U, V. Finally, an obvious analog of Corollary 1.4.5 holds now as well.

We can now tackle the case of coalescing singular values. It is again convenient to first consider the (2×2) case. To begin with, we have the following simple result.

Lemma 2.1.13. *Let $B = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$ be a given matrix. Then, this matrix has two identical nonzero singular values if and only if $B = \sigma Q$, where $Q \in \mathbb{R}^{2\times 2}$ is an orthogonal matrix (and $\sigma \neq 0$).*

Proof. Let $B = U\Sigma V^T$ be the SVD of B. Then, if $\Sigma = \sigma I$, we clearly have $B = \sigma Q$ with $Q = UV^T$. The converse is obvious. □

Naturally, if both singular values are 0, then $B = 0$ in Lemma 2.1.13, which is a trivial multiple of an orthogonal matrix as well, but –as it turns out– the case of both singular values equal to 0 is different and our results do not cover this case, see Remark 2.1.18.

We now have the following result about coalescing singular values.

Theorem 2.1.14 (Signed SVD: (2×2) case)**.** *Consider $B \in \mathcal{C}^e(\Omega, \mathbb{R}^{2\times 2})$, $e \geq 1$. For all $x \in \Omega$, write*

$$B(x) = \begin{bmatrix} a(x) & b(x) \\ c(x) & d(x) \end{bmatrix} ,$$

and let σ_1 and σ_2 be its two continuous singular values, labeled so that $\sigma_1(x) \geq \sigma_2(x) \geq 0$ for all x in Ω. Assume that there exists a unique point $\xi_0 \in \Omega$ where these singular values coincide, $\sigma_1(\xi_0) = \sigma_2(\xi_0)$. Consider the $\mathcal{C}^e$ functions $F, G : \Omega \to \mathbb{R}^2$ given by

$$F(x) = \begin{bmatrix} a^2(x) + c^2(x) - b^2(x) - d^2(x) \\ a(x)b(x) + c(x)d(x) \end{bmatrix}, \; G(x) = \begin{bmatrix} a^2(x) + b^2(x) - c^2(x) - d^2(x) \\ a(x)c(x) + b(x)d(x) \end{bmatrix},$$

and assume that 0 is a regular value for the scalar valued functions given by the 1st and the 2nd components of F and G. Then, consider the $\mathcal{C}^e$ curves Γ_1, Γ_2, Γ_3, and Γ_4, given by the zero-set of the components of F and G: $\Gamma_1 = \{x \in \Omega : a^2(x) + c^2(x) - b^2(x) - d^2(x) = 0\}$, $\Gamma_2 = \{x \in \Omega : a(x)b(x) + c(x)d(x) = 0\}$, and $\Gamma_3 = \{x \in \Omega : a^2(x) + b^2(x) - c^2(x) - d^2(x) = 0\}$, $\Gamma_4 = \{x \in \Omega : a(x)c(x) + b(x)d(x) = 0\}$. Assume that the pair of curves Γ_1 and Γ_2, and also[5] the pair Γ_3 and Γ_4, intersect transversally at ξ_0.

Let Γ be a simple closed curve enclosing the point ξ_0, and let it be parametrized as a $\mathcal{C}^p$ ($p \geq 0$) function γ in the variable t, so that the function $\gamma : t \in \mathbb{R} \to \Omega$ is $\mathcal{C}^p$ and 1-periodic. Let $m = \min(e, p)$, and let B_γ be the $\mathcal{C}_1^m$ function $B(\gamma(t))$, $t \in \mathbb{R}$. Then, for all $t \in \mathbb{R}$, $B_\gamma(t)$ has the signed singular value decomposition

$$B_\gamma(t) = Q_\gamma(t)\Sigma_\gamma(t)Z_\gamma^T(t)$$

such that:

(i) $\Sigma_\gamma \in \mathcal{C}_1^m(\mathbb{R}, \mathbb{R}^{2\times 2})$ and diagonal, $\Sigma_\gamma(t) = \begin{bmatrix} s_1(\gamma(t)) & 0 \\ 0 & s_2(\gamma(t)) \end{bmatrix}$ and $|s_i(\gamma(t))| = \sigma_i(\gamma(t))$, for $i = 1, 2$, and for all $t \in \mathbb{R}$;

(ii) $Q_\gamma, Z_\gamma \in \mathcal{C}_2^m(\mathbb{R}, \mathbb{R}^{2\times 2})$ real orthogonal, and $Q_\gamma(t+1) = -Q_\gamma(t)$, $Z_\gamma(t+1) = -Z_\gamma(t)$, for all $t \in \mathbb{R}$.

[5] *But see Remark 2.1.16*

Proof. Because of Lemma 2.1.13, we have

$$\sigma_1(x) = \sigma_2(x) \iff F(x) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \text{ and } G(x) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

By hypothesis, ξ_0 is the unique root of $F(x)$ and $G(x)$ in Ω, and Γ_1 and Γ_2, and Γ_3 and Γ_4, intersect transversally at ξ_0.

Next, consider the functions B^TB and BB^T. By the assumption on the singular values of B, the eigenvalues of B^TB (which are the same as those of BB^T) in Ω coincide only at ξ_0. Moreover, Theorem 2.1.2 applies to B^TB and BB^T. Let Z_γ and Q_γ be the two orthogonal functions of which in Theorem 2.1.2 relative to B^TB and BB^T, respectively, so that for each $t \in \mathbb{R}$ we have

$$Q^T(\gamma(t))B(\gamma(t))B^T(\gamma(t))Q(\gamma(t)) = \begin{bmatrix} \sigma_1^2(t) & 0 \\ 0 & \sigma_2^2(t) \end{bmatrix}, \quad \text{and}$$

$$Z^T(\gamma(t))B^T(\gamma(t))B(\gamma(t))Z(\gamma(t)) = \begin{bmatrix} \sigma_1^2(t) & 0 \\ 0 & \sigma_2^2(t) \end{bmatrix}, \quad \text{where}$$

$\sigma_1^2(\gamma(t)) > \sigma_2^2(\gamma(t))$, for all $t \in \mathbb{R}$, $Q_\gamma, Z_\gamma \in \mathcal{C}_2^m(\mathbb{R}, \mathbb{R}^{2\times 2})$, and $Q_\gamma(t+1) = -Q_\gamma(t)$ and $Z_\gamma(t+1) = -Z_\gamma(t)$.

Next, consider the function $Q_\gamma^T B_\gamma Z_\gamma$, which is in $\mathcal{C}^m(\mathbb{R}, \mathbb{R}^{2\times 2})$. For all $t \in \mathbb{R}$, write

$$Q^T(\gamma(t))B(\gamma(t))Z(\gamma(t)) = \begin{bmatrix} s_1(\gamma(t)) & z_{12}(\gamma(t)) \\ z_{21}(\gamma(t)) & s_2(\gamma(t)) \end{bmatrix}.$$

Similarly to the proof of Theorem 2.1.11, we must have $z_{12}(\gamma(t)) = 0$, $z_{21}(\gamma(t)) = 0$, for all t. So, we have the decomposition, for all $t \in \mathbb{R}$,

$$Q^T(\gamma(t))B(\gamma(t))Z(\gamma(t)) = \begin{bmatrix} s_1(\gamma(t)) & 0 \\ 0 & s_2(\gamma(t)) \end{bmatrix},$$

where $s_1(\gamma(t)) \neq s_2(\gamma(t))$ for all $t \in \mathbb{R}$, and $|s_1(\gamma(t))| = \sigma_1(\gamma(t)) > |s_2(\gamma(t))| = \sigma_2(\gamma(t))$. Now, for all $t \in \mathbb{R}$, we have $B_\gamma(t+1) = B_\gamma(t)$, $Q_\gamma(t+1) = -Q_\gamma(t)$, $Z_\gamma(t+1) = -Z_\gamma(t)$, and therefore the functions s_1 and s_2 are 1-periodic. □

Remark 2.1.15. The decomposition of B_γ in Theorem 2.1.14 is essentially unique, within the class of $\mathcal{C}^m$ decompositions. The degree of non-uniqueness is given by the ordering of the diagonal and by joint (and global) choices of signs for the columns of Q_γ and Z_γ. In particular, for $\mathcal{C}^m$–decompositions, the statement about periodicity of the functions Q_γ and Z_γ holds unchanged. It is also worth noticing that the functions s_1 and s_2 do not necessarily remain positive along Γ: In fact, if B_γ loses rank, singular value(s) will usually change sign. If we had insisted on having a decomposition with positive singular values, in cases where B_γ lost rank, we would have not been able to retain the $\mathcal{C}^m$ factors Q_γ and Z_γ. Loosely speaking, by insisting on having $\mathcal{C}^m$ orthogonal factors, then the diagonal functions s_1, s_2, must follow their course, and we cannot demand that they are positive in case B_γ loses rank.

Remark 2.1.16. In Theorem 2.1.14, we have assumed that the pair Γ_1, Γ_2, and the pair Γ_3, Γ_4, intersected transversally at ξ_0. In reality, it suffices to assume that one pair of these curves does so, and transversality of the other pair will follow. To verify this statement, we could appeal to Lemma 1.4.9 in a similar way to what we do in Section 2.2.1.2. More directly, a lengthy but otherwise simple computation shows that

$$\det \begin{bmatrix} \nabla\big(a^2(x) + c^2(x) - b^2(x) - d^2(x)\big) \\ \nabla\big(a(x)b(x) + c(x)d(x)\big) \end{bmatrix}_{\xi_0} = \\ = \pm \det \begin{bmatrix} \nabla\big(a^2(x) + b^2(x) - c^2(x) - d^2(x)\big) \\ \nabla\big(a(x)c(x) + b(x)d(x)\big) \end{bmatrix}_{\xi_0}.$$

Remark 2.1.17. The proof of Theorem 2.1.14 begins by observing that the occurrence of a coalescing for a pair of singular values of B can be brought back to the existence of a root for the function F (or equivalently for G). Similarly to Remark 2.1.10, this observation provides a mean of proving that, in general, we are dealing with a codimension 2 phenomenon. In fact, the function F takes values in $\mathbb{R}^2$, and

two constraints need to be satisfied for having a root:

$$\begin{cases} a^2(x) + c^2(x) - b^2(x) - d^2(x) = 0 \\ a(x)b(x) + c(x)d(x) = 0 \end{cases}$$

Without the presence of particular structure in B, these constraints are independent, hence the codimension will actually be 2. In case B possesses some additional structure, the two constrains could be dependent, consequently lowering the codimension of the phenomenon. This is precisely what happens if B is real symmetric. In fact, suppose $b(x) = c(x)$, for all x. Then the two constraints become:

$$\begin{cases} a^2(x) - d^2(x) = 0 \\ b(x)\big(a(x) + d(x)\big) = 0 \end{cases}$$

Clearly, these are not independent, as the sole condition $a(x)+d(x) = 0$ would satisfy both of them. This is the reason why coalescing of singular values has codimension 1 for real symmetric matrices (this has already been pointed out in Remark 1.2.2).

Note that, in case B is real symmetric and positive definite, we have $a(x)+d(x) > 0$ for all x. Consequently, as far as coalescing of singular values is concerned, the phenomenon remains of codimension 2 for real symmetric positive definite matrices.

Remark 2.1.18. We observe that the assumption of transversal intersection of (say) Γ_1 and Γ_2 at the coalescing point ξ_0 rules out the possibility that the singular values be 0 there. This is actually not unexpected, since the request of having a pair of coalescing singular values equal to 0 is a codimension 4 phenomenon: $a = b = c = d = 0$. But, in fact, more is true. The very assumption in Theorem 2.1.14, that 0 be a regular value for the scalar valued functions given by the 1st and the 2nd components of F and G there, implies that we cannot have a pair of coalescing singular values equal to 0, as otherwise there would be no properly defined tangents at all. In other words, our assumptions –here and later on– for coalescing singular values do not allow the coalescing pair of singular values to be 0.

Remark 2.1.19. Similarly to Remark 2.1.4, the assumption of transversality for the curves Γ_1 and Γ_2 (or Γ_3 and Γ_4) at ξ_0 is generic. We will say that ξ_0 is a *generic coalescing point of singular values.*

Using Theorem 2.1.14, and Theorem 2.1.11, we can now tackle the case of a general function in $\mathbb{R}^{n\times n}$, with singular values coalescing at a unique point ξ_0. We first give a definition similar to Definition 2.1.6.

Definition 2.1.20. Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$ have continuous singular values $\sigma_1(x), \ldots,$ $\sigma_n(x)$, $x \in \Omega$, satisfying

$$\sigma_1(x) > \sigma_2(x) > \ldots > \sigma_k(x) \geq \sigma_{k+1}(x) > \ldots > \sigma_n(x) \ , \ \forall x \in \Omega \ ,$$

and

$$\sigma_k(x) = \sigma_{k+1}(x) \Longleftrightarrow x = \xi_0 \in \Omega \ .$$

Let R be a rectangular region $R \subseteq \Omega$ containing ξ_0 in its interior. Moreover, let

(1) $U, V \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$ be $\mathcal{C}^e$ orthogonal functions achieving the reduction of which in Theorem 2.1.11 and Remark 2.1.12-(2):

$$U^T(x)A(x)V(x) = \begin{bmatrix} \Sigma_1(x) & 0 & 0 \\ 0 & B(x) & 0 \\ 0 & 0 & \Sigma_2(x) \end{bmatrix} \ , \ \forall x \in R \ ,$$

where $\Sigma_1 \in \mathcal{C}^e(R, \mathbb{R}^{(k-1)\times(k-1)})$ and $\Sigma_2 \in \mathcal{C}^e(R, \mathbb{R}^{(n-k-1)\times(n-k-1)})$, such that, for all $x \in R$, $\Sigma_1(x) = \mathrm{diag}(\sigma_1(x), \ldots, \sigma_{k-1}(x))$, and $\Sigma_2(x) = \mathrm{diag}(\sigma_{k+2}(x), \ldots,$ $\sigma_n(x))$. Moreover, $B \in \mathcal{C}^e(R, \mathbb{R}^{2\times 2})$ has singular values $\sigma_k(x), \sigma_{k+1}(x)$ for each $x \in R$;

(2) for all $x \in R$, write $B(x) = \begin{bmatrix} a(x) & b(x) \\ c(x) & d(x) \end{bmatrix}$, and define the function F (and G) and the curves Γ_1 and Γ_2 (Γ_3 and Γ_4) as in Theorem 2.1.14, for $x \in R$.

Then, we call ξ_0 a *generic coalescing point of singular values* in Ω, if the curves Γ_1 and Γ_2 (equivalently, Γ_3 and Γ_4) intersect transversally at ξ_0.

It is again true that Definition 2.1.20 holds true regardless of which $\mathcal{C}^e$ transformations U and V we have used to simplify the structure of A. In other words, a result similar to Theorem 2.1.7 holds here as well. In light of Remark 2.1.12-(1), this is a consequence of the following result, whose proof is essentially identical to that of Theorem 2.1.7 and therefore omitted.

Theorem 2.1.21. *Let $B \in \mathcal{C}^e(\Omega, \mathbb{R}^{2\times 2})$ have singular values $\sigma_k(x), \sigma_{k+1}(x)$ for each $x \in \Omega$, and let ξ_0 be the only point where these two singular values coalesce. Write $B(x) = \begin{bmatrix} a(x) & b(x) \\ c(x) & d(x) \end{bmatrix}$, and define the functions F (and G) as in Theorem 2.1.14, satisfying the same assumptions therein, and the curves Γ_1, Γ_2, (and Γ_3, Γ_4), as in Theorem 2.1.14. Let $W, Z \in \mathcal{C}^e(\Omega, \mathbb{R}^{2\times 2})$ be orthogonal, and for all $x \in \Omega$ let*

$$\widetilde{B}(x) = W^T(x)B(x)Z(x) = \begin{bmatrix} \widetilde{a}(x) & \widetilde{b}(x) \\ \widetilde{c}(x) & \widetilde{d}(x) \end{bmatrix} .$$

Define the curves $\widetilde{\Gamma}_1$, $\widetilde{\Gamma}_2$, (and $\widetilde{\Gamma}_3$, $\widetilde{\Gamma}_4$), similarly to how $\Gamma_i, i = 1, ..., 4$, are defined; e.g., $\widetilde{\Gamma}_1 = \{x \in \Omega : \widetilde{a}^2(x) + \widetilde{c}^2(x) - \widetilde{b}^2(x) - \widetilde{d}^2(x) = 0\}$, $\widetilde{\Gamma}_2 = \{x \in \Omega : \widetilde{a}(x)\widetilde{b}(x) + \widetilde{c}(x)\widetilde{d}(x) = 0\}$. Then, if Γ_1 and Γ_2 (equivalently, Γ_3 and Γ_4) intersect transversally at ξ_0, so do $\widetilde{\Gamma}_1$ and $\widetilde{\Gamma}_2$ ($\widetilde{\Gamma}_3$ and $\widetilde{\Gamma}_4$).

We are now ready for the general case of $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$.

Theorem 2.1.22. *Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$. For all $x \in \Omega$, assume that its continuous singular values, $\sigma_1(x), \ldots, \sigma_n(x)$, satisfy*

$$\sigma_1(x) > \sigma_2(x) > \ldots > \sigma_k(x) \geq \sigma_{k+1}(x) > \ldots > \sigma_n(x) \ , \forall x \in \Omega \ ,$$

and

$$\sigma_k(x) = \sigma_{k+1}(x) \Longleftrightarrow x = \xi_0 \ .$$

Let ξ_0 be a generic coalescing point.

Define Γ and γ as in Theorem 2.1.14, and let A_γ be the $\mathcal{C}_1^m$ function $A(\gamma(t))$, $t \in \mathbb{R}$.

Then, for all $t \in \mathbb{R}$, $A_\gamma(t)$ has the (signed) singular value decomposition

$$A_\gamma(t) = U_\gamma(t)\Sigma_\gamma(t)V_\gamma^T(t)$$

satisfying the following conditions:

(i) $\Sigma_\gamma \in \mathcal{C}_1^m(\mathbb{R}, \mathbb{R}^{n\times n})$ and diagonal: $\Sigma_\gamma(t) = \operatorname{diag}(s_1(\gamma(t)), \ldots, s_n(\gamma(t)))$, and $|s_i(\gamma(t))| = \sigma_i(\gamma(t))$, for $i = 1 \ldots, n$, $\forall t \in \mathbb{R}$;

(ii) $U_\gamma, V_\gamma \in \mathcal{C}_2^m(\mathbb{R}, \mathbb{R}^{n\times n})$ real orthogonal, and for all $t \in \mathbb{R}$

$$U_\gamma(t+1) = U_\gamma(t)D \ , \ \ V_\gamma(t+1) = V_\gamma(t)D \ , \quad D = \begin{bmatrix} I_{k-1} & 0 & 0 \\ 0 & -I_2 & 0 \\ 0 & 0 & I_{n-k-1} \end{bmatrix} .$$

Proof. The proof follows similar steps as that of Theorem 2.1.8. First, consider a rectangular region $R \subseteq \Omega$, containing ξ_0. In R, we have a block-SVD

$$\widehat{U}^T(x)A(x)\widehat{V}(x) = \begin{bmatrix} \Sigma_1(x) & 0 & 0 \\ 0 & B(x) & 0 \\ 0 & 0 & \Sigma_2(x) \end{bmatrix} , \forall x \in R ,$$

as in Definition 2.1.20. Then, we take a circle C enclosing ξ_0 and contained in R. Let C be parametrized by a continuous 1-periodic function ρ, and let $B_\rho(t) = B(\rho(t))$, $t \in \mathbb{R}$. Let Q_ρ and Z_ρ be the functions of Theorem 2.1.14 associated to B_ρ. Then, along C, the stated periodicity result holds, since we have

$$U_\rho(t) = \widehat{U}(\rho(t)) \begin{bmatrix} I_{k-1} & 0 & 0 \\ 0 & Q_\rho(t) & 0 \\ 0 & 0 & I_{n-k-1} \end{bmatrix} \qquad \text{and}$$

$$V_\rho(t) = \widehat{V}(\rho(t)) \begin{bmatrix} I_{k-1} & 0 & 0 \\ 0 & Z_\rho(t) & 0 \\ 0 & 0 & I_{n-k-1} \end{bmatrix},$$

and (see Corollary 1.4.5) $\widehat{U}(\rho(t+1)) = \widehat{U}(\rho(t))$ and $\widehat{V}(\rho(t+1)) = \widehat{V}(\rho(t))$ for all t.

Finally, to argue that the same periodicity results hold along Γ, we use the homotopy $h(s,t)$ –which we already used in Theorems 2.1.2 and 2.1.8– to carry U_ρ and V_ρ in U_γ and V_γ. That is, we take the continuous function, which has distinct singular values, $A(h(s,t))$, $(s,t) \in [0,1] \times [0,1]$. Using Theorem 2.1.11, we write $A(h(s,t)) = U(s,t)\Sigma(s,t)V^T(s,t)$, where all factors are continuous, U and V are orthogonal, and Σ is diagonal, $\Sigma(s,t) = \operatorname{diag}\big(s_i(s,t), i = 1,\dots,n\big)$ with $|s_i| = \sigma_i$. Partition $U(s,t)$ and $V(s,t)$ by columns: $U(s,t) = \begin{bmatrix} u_1(s,t) & \cdots & u_n(s,t) \end{bmatrix}$, $V(s,t) = \begin{bmatrix} v_1(s,t) & \cdots & v_n(s,t) \end{bmatrix}$, and consider the functions (for $s \in [0,1]$) $g_j(s) = u_j^T(s,0)u_j(s,1)$ and $f_j(s) = v_j^T(s,0)v_j(s,1)$ for all $j = 1,\dots,n$. Using continuity, we get that $g_j(0) = g_j(1) = f_j(0) = f_j(1) = 1$ for all $j \neq k, k+1$, and $g_j(0) = g_j(1) = f_j(0) = f_j(1) = -1$ for all $j = k, k+1$, and the proof is complete. □

Remark 2.1.23. Similarly to Theorem 2.1.8, also now we have that the $\mathcal{C}^m$ singular vectors of A_γ corresponding to the singular values coalescing at ξ_0 "get upside down" as we complete one loop along the closed curve Γ. Since the singular values of A_γ are distinct, the singular vectors of any $\mathcal{C}^m$ SVD of A_γ cannot increase in period, unless the associated singular values coalesce inside Γ. Finally, we notice that, in Theorem 2.1.22, if the function A_γ had full rank, then we could have obtained that $s_i = \sigma_i$, and positive.

2.2 Generalizations and Main Result

So far, we have studied cases in which a unique pair of eigenvalues of a symmetric function, or singular values of a general function, underwent a generic coalescing at a

point ξ_0. Genericity was related to the way that two smooth curves intersected each other. In this section, we move a step further and consider cases in which several eigenvalues, respectively singular values, coalesce in a region Ω. We stress that, by virtue of the observation made in Remark 1.2.1, eigenvalues (and singular values) will be allowed to coalesce "one pair at a time".

First, we will give –and prove– results for the eigendecomposition of a symmetric matrix valued function. Then, we will state an analogous result for the SVD, but omit the proof since it will be a transparent generalization of the symmetric case.

Before proceeding, we characterize a generic coalescing point, when this is not the only coalescing point in Ω.

Definition 2.2.1. Let $\xi_0 \in \Omega$ be a point where two eigenvalues of the symmetric function $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$, $e \geq 1$, coalesce. Then, we call ξ_0 a *generic coalescing point of eigenvalues* if there exists an open simply connected region $\Omega_0 \subseteq \Omega$, inside which ξ_0 is the unique point where eigenvalues coalesce, and ξ_0 is a generic coalescing point of eigenvalues in Ω_0.

In a nearly identical way, we define a generic coalescing point relatively to the singular values.

Definition 2.2.2. Let $\xi_0 \in \Omega$ be a point where two singular values of the function $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$, $e \geq 1$, coalesce. Then, we call ξ_0 a *generic coalescing point of singular values* if there exists an open simply connected region $\Omega_0 \subseteq \Omega$, inside which ξ_0 is the unique point where singular values of A coalesce, and ξ_0 is a generic coalescing point of singular values in Ω_0.

We are now ready to provide the sought after generalizations. First we will consider the symmetric eigen-problem.

Theorem 2.2.3. *Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$, $e \geq 1$, be symmetric. For all $x \in \Omega$, let $\lambda_1(x), \ldots, \lambda_n(x)$, be its continuous, ordered, eigenvalues, which are distinct except at*

two generic coalescing points ξ_0 *and* ξ_1*. Let* k_1 *and* k_2 *be two distinct indices such that*

$$\lambda_{k_1}(x) = \lambda_{k_1+1}(x) \Longleftrightarrow x = \xi_0 \ , \ \lambda_{k_2}(x) = \lambda_{k_2+1}(x) \Longleftrightarrow x = \xi_1 \ .$$

Without loss of generality, we can take $k_2 > k_1$*.*

Let Γ *be a simple closed curve enclosing the points* ξ_0 *and* ξ_1*, and let it be parametrized as a* $\mathcal{C}^p$ *(*$p \geq 0$*) function* γ *in the variable* t*, so that the function* $\gamma : \ t \in \mathbb{R} \to \Omega$ *is* $\mathcal{C}^p$ *and* 1*-periodic. Let* $m = \min(e, p)$*, and let* A_γ *be the* $\mathcal{C}^m$ *function* $A(\gamma(t))$*,* $t \in \mathbb{R}$*. Then, for all* $t \in \mathbb{R}$*,* $A(\gamma(t))$ *has the eigendecomposition*

$$A_\gamma(t) = U_\gamma(t)\Lambda_\gamma(t)U_\gamma^T(t)$$

such that:

(i) $\Lambda_\gamma \in \mathcal{C}_1^m(\mathbb{R}, \mathbb{R}^{n\times n})$ *and diagonal:* $\Lambda_\gamma(t) = \mathrm{diag}\big(\lambda_1(\gamma(t)), \ldots, \lambda_n(\gamma(t))\big)$*, for all* $t \in \mathbb{R}$*;*

(ii) $U_\gamma \in \mathcal{C}_2^m(\mathbb{R}, \mathbb{R}^{n\times n})$ *real orthogonal, and* $U_\gamma(t+1) = U_\gamma(t)\ D$ *for all* $t \in \mathbb{R}$*, where* $D = D_0D_1$ *with*

(a) $D_0 = \mathrm{diag}(d_j^{(0)},\ j = 1, \ldots, n)$*,* $d_j^{(0)} = 1$*, for* $j \neq k_1, k_1 + 1$*, and* $d_j^{(0)} = -1$*, for* $j = k_1, k_1 + 1$*;*

(b) $D_1 = \mathrm{diag}(d_j^{(1)},\ j = 1, \ldots, n)$*,* $d_j^{(1)} = 1$*, for* $j \neq k_2, k_2 + 1$*, and* $d_j^{(1)} = -1$*, for* $j = k_2, k_2 + 1$*.*

Proof. By assumption, the eigenvalues of A_γ are distinct, and thus the fact that an eigendecomposition of A_γ can be taken of class $\mathcal{C}^m$ is well known (e.g., see [6, 7]). The issue to determine is the periodicity of the eigendecomposition, since the results in [6] do not qualify if the eigendecomposition will be 1-periodic or 2-periodic. We need to look at the behavior of the $\mathcal{C}^m$ eigenvectors of A_γ corresponding to the eigenvalues coalescing inside Γ as we complete a loop along Γ.

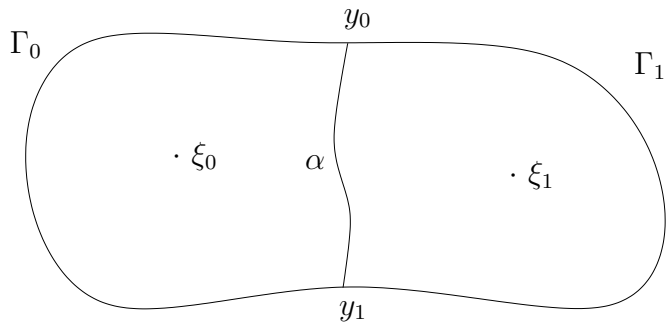

Figure 5: Reference picture for proof of Theorem 2.2.3

Let u_i denote the eigenvector of A_γ corresponding to the eigenvalue λ_i. Let α be a simple curve, which stays inside the region bounded by Γ, connecting two distinct points on Γ, $y_0 = \gamma(t_0)$ to $y_1 = \gamma(t_1)$, with $t_0, t_1 \in [0, 1)$, such that α leaves ξ_0 and ξ_1 on opposite sides (see Figure 5). This is possible, since ξ_0 and ξ_1 are distinct.

With abuse of notation, the symbol $\cup$ will denote the obvious operation of taking the "union" of two curves, whenever this makes sense. Thus, we can let $\Gamma = \Gamma_0 \cup \Gamma_1$, with Γ_0 and Γ_1 having y_0 and y_1 as endpoints. Further, let us require α to be such that $\Gamma_1 \cup \alpha$ is parametrized as a continuous closed curve encircling ξ_1 and $\Gamma_0 \cup (-\alpha)$ as a continuous closed curve encircling ξ_0; here, $-\alpha$ is the same curve transversed in the opposite direction (from y_1 to y_0). Since $\xi_0 \neq \xi_1$, we can always assume that $\Gamma_0 \cup (-\alpha)$ and $\Gamma_1 \cup \alpha$ are enclosed in two domains, $\Omega_0 \subseteq \Omega$ and $\Omega_1 \subseteq \Omega$ respectively, inside which the situation of Definition 2.2.1 applies, relatively to ξ_0 and ξ_1 respectively.

Now, let us consider a $\mathcal{C}^m$ eigendecomposition of A_γ along Γ, starting from and coming back to y_0 (we make the loop only once). Let us denote the eigenvectors of A_γ at the beginning of this loop as u_i^0 and those at the end of the loop as u_i^1, $i = 1, \ldots, n$. We are interested in comparing u_i^0 with u_i^1, for all $i = 1, \ldots, n$. Since the curve α does not contain any coalescing point, the vectors u_i^1's would be the same as if, instead of following the curve Γ, we were to follow the path $\Gamma_0 \cup (-\alpha \cup \alpha) \cup \Gamma_1$. This last path can be interpreted as the union of the two closed loops $\Gamma_0 \cup (-\alpha)$ and $\alpha \cup \Gamma_1$. Let us denote the eigenvectors of A_γ at the end of the first loop as $u_i^{1/2}$, $i = 1, \ldots, n$.

Applying Theorem 2.1.8 to A along the closed curve $\Gamma_0 \cup (-\alpha)$, we conclude that $u_i^0 = -u_i^{1/2}$ for $i = k_1, k_1 + 1$ and $u_i^0 = u_i^{1/2}$ for all other indices. Similarly, we can conclude that $u_i^{1/2} = -u_i^1$ for $i = k_2, k_2+1$ and $u_i^{1/2} = u_i^1$ for all other indices. Putting everything together, we have proven the result. □

Remark 2.2.4. Two special cases are worth being singled out.

(1) In case $k_2 = k_1 + 1$, then from Theorem 2.2.3 we would get

$$D = \begin{bmatrix} I_{k_1-1} & & & & \\ & -1 & & & \\ & & 1 & & \\ & & & -1 & \\ & & & & I_{n-k_1-2} \end{bmatrix}.$$

(2) In case ξ_0 and ξ_1 were generic coalescing points where the same pair of eigenvalues coalesced, the same reasoning used to prove Theorem 2.2.3 would have rendered a 1-periodic, $\mathcal{C}^m$ factor U_γ.

Remark 2.2.4 and the argument used in Theorem 2.2.3 can be extended to prove the following theorem, in which any combination of coalescing of eigenvalues at generic points is allowed.

Theorem 2.2.5. *Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$, $e \geq 1$, be symmetric and let $\lambda_1(x) \geq \ldots \geq \lambda_n(x)$ be its continuous eigenvalues. Suppose that, for every $k = 1, \ldots, n-1$,*

$$\lambda_k(x) = \lambda_{k+1}(x)$$

at d_k distinct generic coalescing points of eigenvalues in Ω, so that there are $\sum_{k=1}^{n-1} d_k$ such points. Let Γ be a simple closed curve enclosing all of these distinct generic coalescing points of eigenvalues. Let Γ be parametrized as a $\mathcal{C}^p$ ($p \geq 0$) function γ in the variable t, so that the function $\gamma : \ t \in \mathbb{R} \to \Omega$ is $\mathcal{C}^p$ and 1-periodic. Let

$m = \min(e,p)$ and $A_\gamma \in \mathcal{C}^m(\mathbb{R}, \mathbb{R}^{n\times n})$ defined as $A_\gamma(t) = A(\gamma(t))$, for all $t \in \mathbb{R}$. Then, for all $t \in \mathbb{R}$,

$$A_\gamma(t) = U_\gamma(t)\Lambda_\gamma(t)U_\gamma^T(t)$$

such that:

(i) $\Lambda_\gamma \in \mathcal{C}_1^m(\mathbb{R}, \mathbb{R}^{n\times n})$ is diagonal: $\Lambda_\gamma(t) = \operatorname{diag}\big(\lambda_1(\gamma(t)), \ldots, \lambda_n(\gamma(t))\big)$, for all $t \in \mathbb{R}$;

(ii) $U_\gamma \in \mathcal{C}^m(\mathbb{R}, \mathbb{R}^{n\times n})$ orthogonal, with

$$U_\gamma(t+1) = U_\gamma(t)\ D\ ,\ \forall t \in \mathbb{R}\ ,$$

where D is a diagonal matrix of ± 1 given as follows:

$$D_{11} = (-1)^{d_1},\quad D_{kk} = (-1)^{d_{k-1}+d_k} \text{ for } k = 2, \ldots, n-1,\quad D_{nn} = (-1)^{d_{n-1}}\ .$$

In particular, if $D = I_n$, then U_γ is 1-periodic, otherwise U_γ is 2-periodic.

Proof. The proof is by induction on the number of coalescing points. We know that the result is true for 1 and 2 coalescing points. Thus, we will assume the result to be true for $N-1$ distinct generic coalescing points, and will show it for N distinct generic coalescing points, where $N = \sum_{k=1}^{n-1} d_k$.

Since the coalescing points are distinct, we can always separate one of them, call it ξ_N, from the other $N-1$, with a curve α not containing coalescing points, and which stays inside the region bounded by Γ, joining two distinct points on Γ, y_0 and y_1 (see Figure 6). Let $j,\ 1 \le j \le n-1$, be the index for which $\lambda_j(\xi_N) = \lambda_{j+1}(\xi_N)$.

Now, with a construction identical to the one we used in the proof of Theorem 2.2.3, and with the same notation we used there, consider a smooth eigendecomposition of A_γ along Γ, starting from and coming back to y_0 (the loop is done once). Denote the matrix of eigenvectors of A_γ at the beginning of this loop as U_0 and that at the end of the loop as U_1. Since the curve α does not contain any coalescing point,

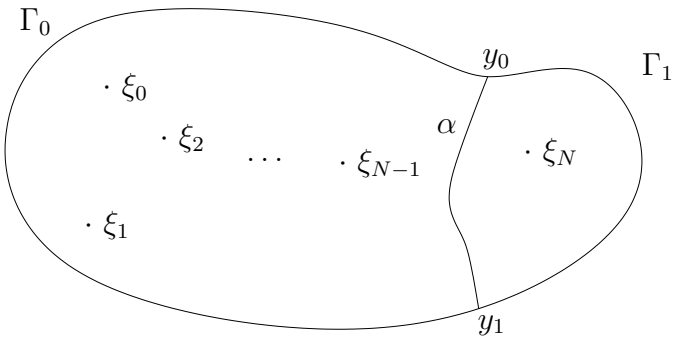

Figure 6: Reference picture for proof of Theorem 2.2.5

the matrix U_1 would be the same as if, instead of following the curve Γ, we were to follow the path $(\Gamma_0 \cup (-\alpha)) \cup (\alpha \cup \Gamma_1)$. Denote the matrix of eigenvectors of A_γ at the end of the first loop by $U_{1/2}$. Using the induction hypothesis along the closed curve $\Gamma_0 \cup (-\alpha)$, we have

$$U_0 = U_{1/2}\widehat{D}\,,$$

where $\widehat{D}$ is a diagonal matrix $\widehat{D} = \mathrm{diag}(\widehat{d}_1, \ldots \widehat{d}_n)$, with $\widehat{d}_k = d_k$, for all $k \neq j$, and $\widehat{d}_j = d_j - 1$.

By looking at what happens on the second loop, by virtue of Theorem 2.1.8, we have that all columns of $U_{1/2}$ coincide with those of U_1, except for the j-th and $(j+1)$-st ones which have changed in sign. Putting everything together, we have $U_0 = U_1 D$ with D as given in the statement of the Theorem. □

A result similar to Theorem 2.2.5 holds for the singular value decomposition of a general function in a domain where singular values coalesce at generic coalescing points of singular values. The precise statement follows without proof, since the proof is a clear adaptation of the previous ones.

Theorem 2.2.6. *Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$, $e \geq 1$. Let $\sigma_1(x) \geq \ldots \geq \sigma_n(x)$ be its continuous singular values and suppose that, for every $k = 1, \ldots, n-1$,*

$$\sigma_k(x) = \sigma_{k+1}(x)$$

at d_k distinct generic coalescing points of singular values in Ω. Let Γ be a simple closed curve enclosing all these coalescing points, parametrized by the 1-periodic function $\gamma \in \mathcal{C}_1^p(\mathbb{R}, \Omega)$. Let $m = \min(e, p)$ and $A_\gamma \in \mathcal{C}^m(\mathbb{R}, \mathbb{R}^{n\times n})$ defined as $A_\gamma(t) = A(\gamma(t))$, for all $t \in \mathbb{R}$. Then, for all $t \in \mathbb{R}$,

$$A_\gamma(t) = U_\gamma(t) S_\gamma(t) V_\gamma^T(t)$$

such that:

(i) $S_\gamma \in \mathcal{C}_1^m(\mathbb{R}, \mathbb{R}^{n\times n})$ diagonal: $\Sigma_\gamma(t) = \operatorname{diag}\big(s_1(\gamma(t)), \ldots, s_n(\gamma(t))\big)$, for all $t \in \mathbb{R}$, and $|s_i(\gamma(t))| = \sigma_i(\gamma(t))$, for all $i = 1, \ldots, n$, and all $t \in \mathbb{R}$;

(ii) $U_\gamma, V_\gamma \in \mathcal{C}^m(\mathbb{R}, \mathbb{R}^{n\times n})$ orthogonal, with

$$U_\gamma(t+1) = U_\gamma(t)\ D\ ,\ V_\gamma(t+1) = V_\gamma(t)\ D\ ,\quad \forall t \in \mathbb{R}\ ,$$

where D is as in Theorem 2.2.5.

2.2.1 Lack of Transversality and Multiplicity of Coalescing

At this point, we observe that our results have been "built upon" Theorem 2.1.2. The assumptions in that theorem were motivated by the realization that when the curves $\Gamma_1 = \{x : (a-d)(x) = 0\}$ and $\Gamma_2 = \{x : b(x) = 0\}$ intersect each other, they are expected to do so transversally. This is the generic case. If not, we would need to have three conditions satisfied at a point ξ_0 where eigenvalues coalesce: $(a-d)(\xi_0) = 0$, $b(\xi_0) = 0$, and $\det \begin{bmatrix} \nabla(a-d) \\ \nabla b \end{bmatrix}_{\xi_0} = 0$. This means that non-transversal intersection has codimension 3 and thus is a nongeneric property for functions of 2 parameters.

However, closer inspection of the proof of Theorem 2.1.2 reveals that transversal intersection of Γ_1 and Γ_2 is not strictly necessary. What we really need in Theorem 2.1.2 is that the curves Γ_1 and Γ_2 *cross each other at* ξ_0. Let us clarify what we mean by this.

2.2.1.1 Crossing: Symmetric Eigenproblem

First, consider the case of symmetric $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{2\times 2})$, $e \geq 1$. Notation is the same as in Theorem 2.1.2.

As usual, we assume that 0 be a regular value for the functions $a - d$ and b, so that the curves $\Gamma_{1,2}$ through ξ_0 are $\mathcal{C}^e$ curves, with non-vanishing gradients along the curves. Consider the two vectors $\nabla(a-d)(\xi_0)$ and $\nabla b(\xi_0)$, both of which are non zero. If these vectors are independent, then the two curves intersect transversally in the sense of Definition 1.5.3. If they are not independent, then we must have $\nabla(a-d)(\xi_0) = \kappa\nabla b(\xi_0)$ for some constant $\kappa \neq 0$. In particular, the first or the second component of these two vectors, or both, must be different from 0. By virtue of the implicit function theorem, this means that –locally, near ξ_0– both Γ_1 and Γ_2 may be parametrized in terms of the same variable: x_1 if the second component of the tangent vectors is not 0, or x_2 if the first is not 0. Without loss of generality, we can assume that $(a-d)_{x_2}(\xi_0) \neq 0 \neq b_{x_2}(\xi_0)$. So, near ξ_0, $\Gamma_1 = \{x = \big(t,\ g_1(t)\big)\}$ and $\Gamma_2 = \{x = \big(t,\ g_2(t)\big)\}$, for t in a neighborhood of the origin, say $t \in (-\eta,\ \eta)$, $\eta > 0$. By construction, $g_1(0) = g_2(0)$. Define $g(t) = g_1(t) - g_2(t)$, $t \in [-\eta, \eta]$. Again appealing to the implicit function theorem, g_1, g_2, and hence g, are $\mathcal{C}^e$ functions of t. Also, notice that $g_1'(t) = -(a-d)_{x_1}/(a-d)_{x_2}$, at points $x = \xi_0 + (t, g_1(t))$, and $g_2'(t) = -b_{x_1}/b_{x_2}$, at points $x = \xi_0 + (t, g_2(t))$, $t \in (-\eta, \eta)$. In particular, since $\nabla(a-d)(\xi_0) = \kappa\nabla b(\xi_0)$, we have $g'(0) = 0$.

Definition 2.2.7. We say that Γ_1 and Γ_2 *cross each other at* ξ_0 if they intersect transversally in the sense of Definition 1.5.3, or the function g above changes sign as t goes through 0.

We observe that, by using the above characterization of curves crossing, instead of their transversal intersection, the construction and proof of Theorem 2.1.2 hold unchanged, as it is plainly seen.

Now, assuming enough differentiability for the function A, and hence for g, one can characterize more precisely the order of contact of the curves Γ_1 and Γ_2. In the discussion that follows, the degree of differentiability e of $A \in \mathcal{C}^e$, will be assumed to be sufficiently high so that all derivative terms we take make sense. We reiterate that $g(0) = 0 = g'(0)$ in case Γ_1 and Γ_2 do not intersect transversally.

Definition 2.2.8 (Order of Contact)**.** Γ_1 and Γ_2 have a contact of order $p = 0$ at ξ_0 if they intersect transversally there. Also, Γ_1 and Γ_2 have a contact of order $p \geq 1$ at ξ_0 if $g^{(j)}(0) = 0$, $0 \leq j \leq p$, but $g^{(p+1)}(0) \neq 0$.

Definition 2.2.8 allows us to formally define multiplicity of coalescing for the eigenvalues themselves, by tracing it back to the order of contact of the curves Γ_1 and Γ_2 at ξ_0. More precisely, we have:

Definition 2.2.9 (Multiplicity of Coalescing of Eigenvalues)**.** We say that the eigenvalues have a coalescing of multiplicity $p \geq 1$ at ξ_0 if the curves Γ_1 and Γ_2 have a contact of order $p - 1$ at ξ_0, in the sense of Definition 2.2.8.

Remark 2.2.10 (Crossing Equals Even Order of Contact)**.** According to Definitions 2.2.7 and 2.2.9, Γ_1 and Γ_2 cross each other at ξ_0 if there exist an index $k \geq 1$ for which $g^{(j)}(0) = 0$, $0 \leq j \leq 2k$, but $g^{(2k+1)}(0) \neq 0$. In other words, crossing is the same as even order of contact.

Next, we want to extend the above concepts to the case of a symmetric function $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$. Informally, we want to apply the (2×2) case locally. We have the following immediate adaptation of Definitions 2.1.6 and 2.2.1, with the same notation used there.

Definition 2.2.11. Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$ be a symmetric function with ordered eigenvalues $\lambda_1(x) \geq \lambda_2(x) \geq \cdots \geq \lambda_n(x)$, for all $x \in \Omega$. Let $\xi_0 \in \Omega$ be a coalescing point, that is, for some k, $1 \leq k \leq n$, $\lambda_k(\xi_0) = \lambda_{k+1}(\xi_0)$. Let R be a sufficiently small

rectangular region $R \subseteq \Omega$ containing ξ_0 in its interior, and such that ξ_0 is the only coalescing point in R. Let $Q \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$ be a $\mathcal{C}^e$ orthogonal function achieving the reduction:

$$Q^T(x)A(x)Q(x) = \begin{bmatrix} \Lambda_1(x) & 0 & 0 \\ 0 & P(x) & 0 \\ 0 & 0 & \Lambda_2(x) \end{bmatrix} , \forall x \in R ,$$

where $P \in \mathcal{C}^e(R, \mathbb{R}^{2\times 2})$ is symmetric with eigenvalues $\lambda_k(x), \lambda_{k+1}(x)$, and Λ_1 and Λ_2 are diagonal functions containing the other eigenvalues. Let $P(x) = \begin{bmatrix} a(x) & b(x) \\ b(x) & d(x) \end{bmatrix}$, and let the curves Γ_1 and Γ_2 be as in the (2×2) case. Then, we say that the curves Γ_1 and Γ_2 cross each other at ξ_0 if they do so according to Definition 2.2.7. Finally, with the obvious extension of Definitions 2.2.8 and 2.2.9 to this case, we will say that eigenvalues have a coalescing of multiplicity $p \geq 1$ at ξ_0 if the curves Γ_1 and Γ_2 have a contact of order $p - 1$ there.

To justify Definition 2.2.11, we need to show that –in this $(n \times n)$ case– the above concept of order of contact is independent of the choice of $\mathcal{C}^e$ transformation Q which achieves the block-diagonalization of Definition 2.2.11. This is actually correct, see below. First of all, because of Theorem 2.1.7, we know that transversal intersection (hence, contact of order 0) is independent of the choice of the $\mathcal{C}^e$ transformation Q. In general, we have the following result.

Theorem 2.2.12. *Let A, ξ_0, R, Q and P be as in Definition 2.2.11. Let $U \in \mathcal{C}^e(R, \mathbb{R}^{n\times n})$ be another orthogonal function achieving a similar block reduction:*

$$U^T(x)A(x)U(x) = \begin{bmatrix} \Lambda_1(x) & 0 & 0 \\ 0 & \widetilde{P}(x) & 0 \\ 0 & 0 & \Lambda_2(x) \end{bmatrix} , \forall x \in R .$$

Write $\widetilde{P}(x) = \begin{bmatrix} \widetilde{a}(x) & \widetilde{b}(x) \\ \widetilde{b}(x) & \widetilde{d}(x) \end{bmatrix}$. Then, if the curves $\Gamma_1 = \{x \in R : (a - d)(x) = 0\}$

and $\Gamma_2 = \{x \in R :\ b(x) = 0\}$ *have a contact of order* $p \geq 1$ *at* ξ_0*, so do the curves* $\widetilde{\Gamma}_1 = \{x \in R :\ (\widetilde{a} - \widetilde{d})(x) = 0\}$ *and* $\widetilde{\Gamma}_2 = \{x \in R :\ \widetilde{b}(x) = 0\}$.

Proof. We begin by observing that, from (4-5) in Theorem 2.1.7, it follows that $\nabla(a - d)$, ∇b, $\nabla(\widetilde{a} - \widetilde{d})$, $\nabla\widetilde{b}$ are all parallel to each other in case Γ_1 and Γ_2 have a non-transversal intersection at ξ_0. Therefore, by the implicit function theorem, near ξ_0 the curves Γ_1, Γ_2, $\widetilde{\Gamma}_1$, $\widetilde{\Gamma}_2$ can be parametrized in terms of the same variable. So, without loss of generality, we can assume that they are parametrized in terms of x_1: $\Gamma_1 = \{x = \big(t,\ g_1(t)\big)\}$, $\Gamma_2 = \{x = \big(t,\ g_2(t)\big)\}$, $\widetilde{\Gamma}_1 = \{x = \big(t,\ \widetilde{g}_1(t)\big)\}$, $\widetilde{\Gamma}_2 = \{x = \big(t,\ \widetilde{g}_2(t)\big)\}$, for $t \in (-\eta,\ \eta)$, $\eta > 0$. Let $g = g_1 - g_2$, $\widetilde{g} = \widetilde{g}_1 - \widetilde{g}_2$. By hypothesis, $t = 0$ is a zero of multiplicity p for g. We want to show that the same is true for $\widetilde{g}$. By virtue of Theorem 2.1.20, ξ_0 must be a point of non-transversal intersection also for $\widetilde{\Gamma}_1$ and $\widetilde{\Gamma}_2$, so that $\widetilde{g}(0) = \widetilde{g}'(0) = 0$. Let us suppose, by contradiction, that $t = 0$ is a zero of multiplicity q, $q < p$, for $\widetilde{g}$. Given that Γ_1 and Γ_2 intersect at ξ_0 with order of contact p, it is possible to additively perturb $P(x)$ with a symmetric matrix function $\varepsilon E(x)$, $\|E\| = 1$, so that $P(x) + \varepsilon E(x)$ has p distinct coalescing points in a small neighborhood of ξ_0, for all $\varepsilon \neq 0$ sufficiently small, say $0 < \varepsilon < \bar{\varepsilon}$. If $\bar{\varepsilon}$ is chosen small enough, the perturbation will modify in an obvious way g and $\widetilde{g}$ into, respectively, g_ε and $\widetilde{g}_\varepsilon$, both functions being defined in $[-\eta, \eta]$. Now, the zeros of g_ε and $\widetilde{g}_\varepsilon$ are coalescing points for the eigenvalues, and since P and $\widetilde{P}$ are similar, g and $\widetilde{g}$ have exactly the same p distinct zeros in $[-\eta, \eta]$, for all $\varepsilon > 0$, all approaching 0 as ε goes to 0. Therefore, in the limit as ε approaches 0, we can conclude that $\widetilde{g}$ has a zero of multiplicity $p > q$ for $t = 0$, which contradicts our earlier assumption. The case $q > p$ needs not be discussed, as the roles of g and $\widetilde{g}$ can be interchanged. □

2.2.1.2 Crossing: SVD

We need to extend the results relative to non-generic crossing we saw for the symmetric eigenproblem to the case of singular values. Recall that, see Remark 2.1.18, the

pair of singular values which will coalesce cannot be 0. Thus, there is a neighborhood of the coalescing singular values where the function A is invertible. Now, appealing to Lemma 1.4.9, we can write $A = QP$, with Q orthogonal, P symmetric and positive definite, and both as smooth as A in a neighborhood of a coalescing point. Obviously, the singular values are the same as the eigenvalues of P, and the functions AA^T and A^TA are smoothly similar via Q. As a consequence, we can immediately extend Definitions 2.2.7, 2.2.8, 2.2.9, and 2.2.11, to the case of singular values.

Finally, it is simple to appreciate that, with the obvious modifications for the singular values, all of Theorems 2.1.8, 2.1.14 and 2.1.22 hold unchanged in their proofs and conclusions by replacing the assumption of generic coalescing point with that of coalescing point of odd multiplicity for the eigenvalues (singular values), as appropriate. Likewise, by replacing in Definitions 2.2.1 (and similar modifications in 2.2.2) the request that ξ_0 be a "generic coalescing point" with that of "coalescing point of odd multiplicity for the eigenvalues (singular values)", also Theorems 2.2.3, 2.2.5, and 2.2.6, hold unchanged.

2.2.2 Main Result

The last part of this Section is devoted to our main Theorem.

The considerations of Section 2.2.1 extend Theorems 2.2.5 and 2.2.6 to the case of coalescing points of multiplicity higher than 1. With the aid of Lemma 1.4.8, which excludes the coexistence of continuous decompositions of period 1 and 2 for the same periodic matrix function, we can conclude that the reverse of all the Theorems of this chapter holds true: If some eigenvectors (respectively, singular vectors) of a smooth matrix function A change sign under continuation around a closed loop Γ, the relevant eigenvalues (respectively, singular values) must have coalesced somewhere inside Γ.

This is precisely the content of the following result, which forms the basis for algorithms (see Chapter 4) which attempt to locate coalescing points.

Theorem 2.2.13. *Let $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$, $e \geq 1$, symmetric, with continuous and ordered eigenvalues (respectively, $A \in \mathcal{C}^e(\Omega, \mathbb{R}^{n\times n})$ have continuous and ordered singular values). Let Γ be a simple closed curve in Ω, with no coalescing point for the eigenvalues (respectively, singular values) of A on it. Let γ be a $\mathcal{C}^p$, 1-periodic parametrization for Γ, let $m = \min(e, p)$, and let $A_\gamma \in \mathcal{C}^m(\mathbb{R}, \mathbb{R}^{n\times n})$, U_γ (and V_γ) be defined as in Theorems 2.2.5 (and 2.2.6). Finally, let $U_0 = U_\gamma(0)$ and $U_1 = U_\gamma(1)$, and define $D = U_0^T U_1$.*

Next, let $2q$ be the (even) number of indices k_i, $k_1 < k_2 < \cdots < k_{2q}$, for which $D_{k_i k_i} = -1$. Let us group these indices in pairs $(k_1, k_2), \ldots, (k_{2q-1}, k_{2q})$. Then, λ_k and λ_{k+1} (respectively, σ_k and σ_{k+1}) coalesce an odd number of times, counting multiplicities, or infinitely many times, inside the region encircled by Γ, if $k_{2j-1} \leq k < k_{2j}$ for some $j = 1, \ldots, q$.

Remarks 2.2.14. Theorem 2.2.13 is an immediate consequence of Theorems 2.2.5 and 2.2.6, and Lemma 1.4.8, therefore its proof is omitted. Nonetheless, some points need to be clarified.

(a) In the statement, we claim that the number of indices k_i for which $D_{k_i k_i} = -1$ is even. Here we justify the claim only for the eigenproblem, the SVD case being similar. Let A, Γ, γ, U_γ, Λ_γ and D be as usual. Then we have $A_\gamma(t) = U_\gamma(t)\Lambda_\gamma(t)U_\gamma^T(t)$, with $U_\gamma(t+1) = U_\gamma(t)\, D$. Being $U_\gamma(t)$ continuous for all t, we have that its determinant must be constantly equal either to 1 or to -1. Now, since $D = U_\gamma^T(0)U_\gamma(1)$, we have that $\det(D) = 1$, hence the claim is justified.

(b) In the conclusion of Theorem 2.2.13, we cannot rule out the possibility that the eigenvalues (singular values) coalesce infinitely many times, though this is certainly not a generic situation. At the same time, the case of infinitely many coalescing points may as well go undetected. To clarify, see Example 2.2.15 below.

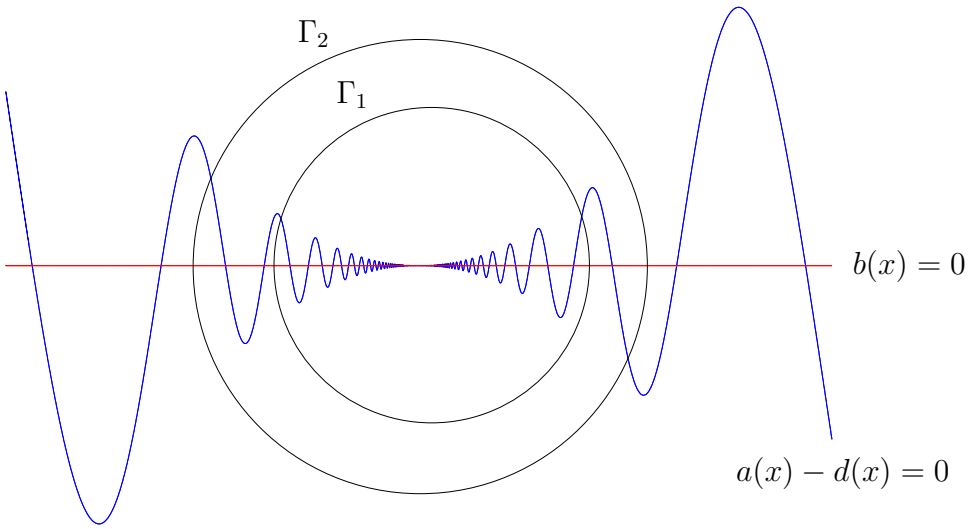

Figure 7: Infinitely many intersections: period 1 or 2.

(c) We stress that we expect all coalescing points to have multiplicity 1 (this is the generic property). In this case, we can illustrate Theorem 2.2.13 as follows. Suppose we have $n \geq 4$, and D with $D_{11} = -1 = D_{44}$, all other D_{ii}'s being 1. Then, we expect that inside the region encircled by Γ, the pairs (λ_1, λ_2), (λ_2, λ_3), and (λ_3, λ_4), have coalesced.

Example 2.2.15. Take $x = (x_1, x_2)$, and consider the $\mathcal{C}^p$ function ($p \geq 1$)

$$A(x) = \begin{bmatrix} x_2 & x_2 \\ x_2 & x_1^{p+1} \sin(1/x_1) \end{bmatrix},$$

whose eigenvalues coalesce on the x_1-axis (and only there), infinitely many times. Let Γ be a simple closed curve containing the origin inside. From the considerations in the proof of Theorem 2.1.2, it follows that the eigenvector corresponding to the largest eigenvalue of A crosses the x_1-axis (respectively, x_2-axis) exactly when Γ crosses the curve $b(x) = 0$ and $a(x) - d(x) > 0$ (respectively, $a(x) - d(x) < 0$), see Figure 7. Therefore, with usual notation, we may have that U_γ is 1-periodic: $U_1 = U_0$ (e.g., along Γ_1) or 2-periodic: $U_1 = -U_0$ (e.g., along Γ_2).

Remark 2.2.16. Unfortunately, it is impossible to distinguish, by the arguments above, whether some pair of eigenvalues (respectively, singular values) coalesce an even number of times or do not coalesce at all inside Γ.

CHAPTER III

ANALYSIS OF THE EFFECT OF PERTURBATIONS

In Section 1.2.1 we have discussed genericity and codimension of coalescing of singular values (and eigenvalues). Therein, we have observed that, for a smooth matrix valued function depending on one parameter, having a pair of coalescing singular values is not a generic property (and hence it should not be observed in 1-parameter functions). But, it is a generic property for smooth functions depending on two parameters, in which case the generic property is that singular values will coalesce at isolated values of the parameters.

Studying the way degenerate cases break down under perturbation turns out to be quite insightful, and understanding this situation is also of help in understanding how the algorithms of Chapter 4 work. For these reasons, in this Chapter we look at specific examples that illustrate these situations. We restrict to (2×2) cases, so to keep the arguments simple without losing any essential feature. Also, we will consider only the case of singular values; similar considerations apply to the case of eigenvalues of symmetric matrices.

3.1 One Parameter Case

We illustrate this case with the following Example.

Example 3.1.1. Let $t \in [0, 1]$,

$$\Sigma(t) = \begin{bmatrix} 1 + (t - .5)^2 & 0 \\ 0 & 1.125 \end{bmatrix}, \; U(t) = \begin{bmatrix} \cos(t) & \sin(t) \\ \sin(t) & -\cos(t) \end{bmatrix}$$

and $A(t) = U(t)\,\Sigma(t)\,U^T(t)$. The singular values of A coalesce for two values of the parameter t, see Figure 8. Note that, in order to preserve the smoothness of the SVD

factors, singular values have been allowed to exchange ordering. Let us perturb A by adding to it a (2×2) matrix E, with entries chosen at random (uniformly distributed in $[-1, 1]$) and of small norm:

$$E = 10^{-2} \begin{bmatrix} 0.843... & -0.647... \\ 0.476... & -0.188... \end{bmatrix}.$$

Figure 9 shows the effect of the perturbation on the singular values and singular vectors. As we expected, the singular values of $A + E$ do not coalesce. Actually, they come very close to coalescing, but then they suddenly "curve away". This is the classical "veering" (also known as "avoided crossing") phenomenon, of which we have already reported in Section 1.3.2. We stress once again that this phenomenon is not an anomaly, but rather a typical behavior for 1-parameter dependent matrix functions. Something very interesting occurs concerning the singular vectors: Their entries seem to develop nearly jump discontinuities in correspondence of the points where the singular values of the unperturbed matrix A coalesce. This is also what one should expect to observe in general. But, while the "veering" effect is easily justified by a codimension argument, the "jump" behavior of the singular vectors deserves more explanations. In Example 3.1.3 below we show that it is a side-effect of the breakdown of a coalescing pair of singular values, essentially due to a sharp 90° rotation of the corresponding singular vectors.

Remark 3.1.2. The results whose outcome we reported in Figures 8 and 9 have been obtained with the 1-d solver which computes a smooth path of SVD decompositions; see Section 4.1.2. It is noteworthy to point out that the computations for the perturbed case are far more expensive and delicate than those for the unperturbed case (at first, this appears to be almost counterintuitive!). The reason is that the closer one gets to the coalescing point of the unperturbed problem, the harder one has to work to maintain smooth decompositions for the orthogonal factors. This is

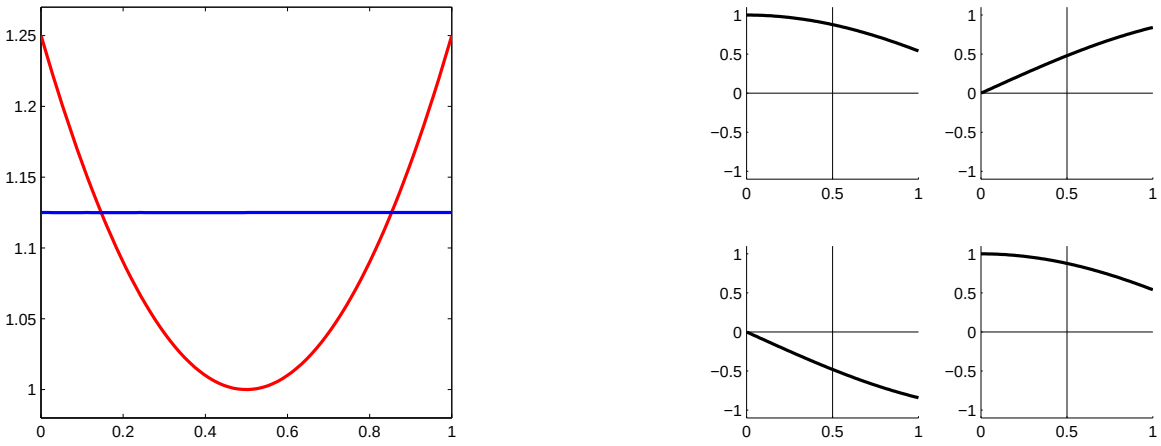

Figure 8: Singular values of A (left) and each of the entries of one matrix of singular vectors (right).

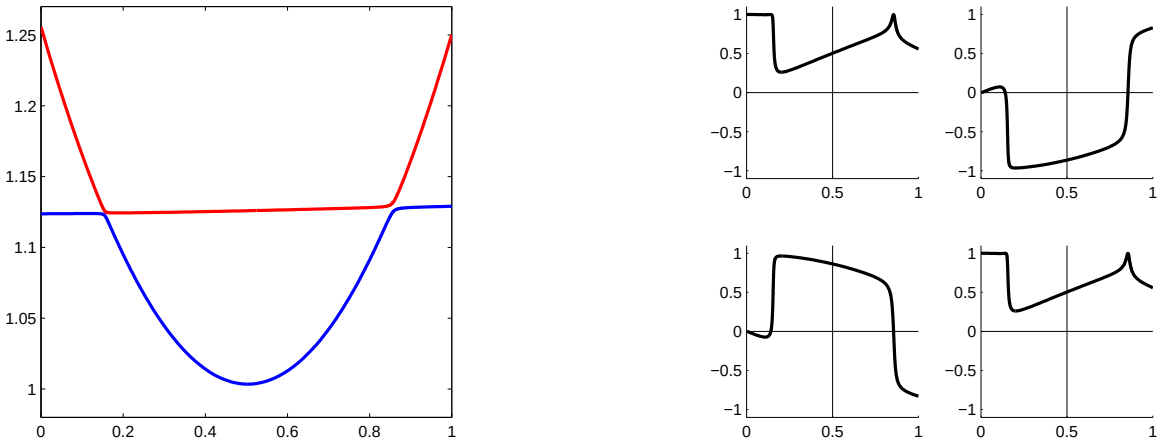

Figure 9: Singular values of $A + E$ (left) and each of the entries of one matrix of singular vectors (right).

due to the high speed of rotation of these factors near the coalescing point of the unperturbed problem. To witness, to complete the smooth path of SVDs for the two cases of Figures 8 and 9 required 85, respectively 670, continuation steps.

Example 3.1.3 (Jump of Singular Vectors)**.** Let $A \in \mathcal{C}^e(\mathbb{R}, \mathbb{R}^{2\times 2})$, and let its singular values σ_1, σ_2, and corresponding singular vectors u_1, u_2, v_1, v_2, be all smooth functions. Suppose $\sigma_1(t) = \sigma_2(t)$ if and only if $t = \bar{t}$. Let $E \in \mathcal{C}^e(\mathbb{R}, \mathbb{R}^{2\times 2})$ be of small norm and such that $E(t) = 0$ for all $|t - \bar{t}| > \delta$, for some small $\delta > 0$. We will think of E as a (compactly supported) perturbation function. Let $\widetilde{\sigma}_1$ and $\widetilde{\sigma}_2$ be the smooth singular values of the "perturbed" function $A + E$ and $\widetilde{u}_1$, $\widetilde{u}_2, \widetilde{v}_1$, $\widetilde{v}_2$, the corresponding smooth singular vectors. Now assume that $\widetilde{\sigma}_1(t) \neq \widetilde{\sigma}_2(t)$ for all t (in particular, near $t = \bar{t}$). In other words, we are assuming that the perturbation actually breaks down the coalescing. This last assumption is not at all restrictive, since we expect that just about every E will achieve it. The argument that follows is based on simple observations on the degree of uniqueness of the SVD of a matrix with distinct singular values. Since $A(t) = A(t) + E(t)$ for all $t < \bar{t} - \delta$, for the same values of t we have that $\sigma_1(t) = \widetilde{\sigma}_1(t)$ and $\sigma_2(t) = \widetilde{\sigma}_2(t)$. The degree of non-uniqueness of the corresponding singular vectors is given by their signs. Therefore we can, and do, assume that $u_1(t) = \widetilde{u}_1(t)$, $u_2(t) = \widetilde{u}_2(t)$, for $t < \bar{t} - \delta$, and the same thing for the right singular vectors (v_1, v_2 and $\widetilde{v}_1, \widetilde{v}_2$). For $t = \bar{t}$, $\sigma_1(t)$ and $\sigma_2(t)$ coalesce and cross each other, while $\widetilde{\sigma}_1(t)$ and $\widetilde{\sigma}_2(t)$ do not coalesce for any value of t, see Figure 10. On the other hand, we have $A(t) = A(t) + E(t)$ for all $t > \bar{t} + \delta$. Hence, for the same values of t, we must have $\sigma_1(t) = \widetilde{\sigma}_2(t)$, $\sigma_2(t) = \widetilde{\sigma}_1(t)$, $u_1(t) = \pm\widetilde{u}_2(t)$, $u_2(t) = \pm\widetilde{u}_1(t)$, for $t > \bar{t} + \delta$, and the same thing for the right singular vectors. In other words, the singular vector $\widetilde{u}_1$ (resp. $\widetilde{u}_2$), which starts off aligned to u_1 (resp. u_2), rotates toward the direction of u_2 (resp. u_1) over the interval $[\bar{t} - \delta, \bar{t} + \delta]$. Since $\begin{bmatrix} \widetilde{u}_1 & \widetilde{u}_2 \end{bmatrix}$ is a smooth

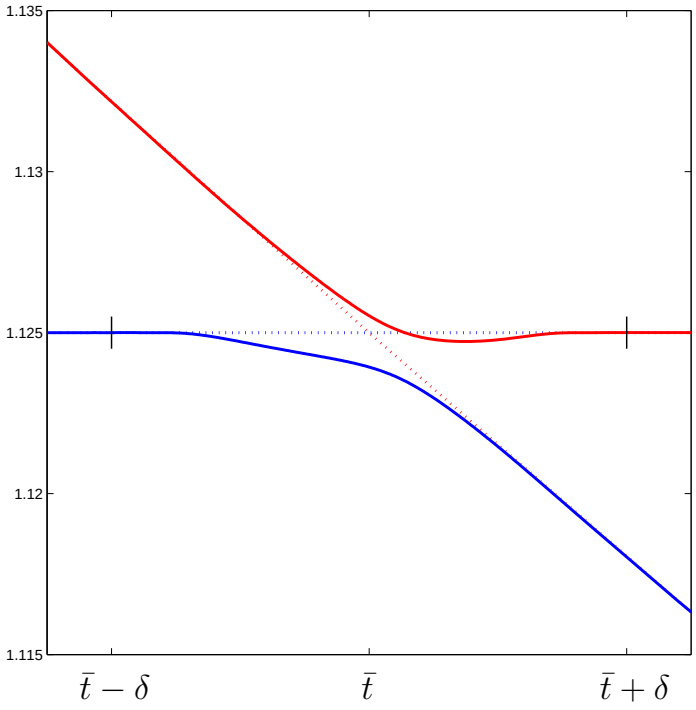

Figure 10: Dotted lines represent σ_1 (red) and σ_2 (blue), solid lines represent $\widetilde{\sigma}_1$ (red) and $\widetilde{\sigma}_2$ (blue); the Figure also shows the support of the perturbation

orthonormal frame for all $t \in \mathbb{R}$, the only (two) possibilities, for $t > \bar{t} + \delta$, are:

$$\begin{bmatrix} \widetilde{u}_1 & \widetilde{u}_2 \end{bmatrix} = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \begin{bmatrix} 0 & \pm 1 \\ \mp 1 & 0 \end{bmatrix}.$$

The same thing can be argued concerning the right singular vectors. This justifies our claim: If a 1-parameter matrix function is perturbed so that a coalescing of singular values breaks down, the corresponding singular vectors will perform a sharp $\pm 90°$ rotation. Note that the rotation occurs over an interval of length 2δ, where $\delta > 0$ is arbitrarily small. Therefore, in principle, the rotation may be made as sharp as desired.

Remark 3.1.4. In applications, one often knows the value of the function $A(\cdot)$ only at a discrete set of points $t_0, \ldots, t_N$, and it is natural to try to find an approximation to the function A via cubic splines or some other interpolation tool. More in general, it is easy to imagine situations where A is known only up to a certain error. Example 3.1.1

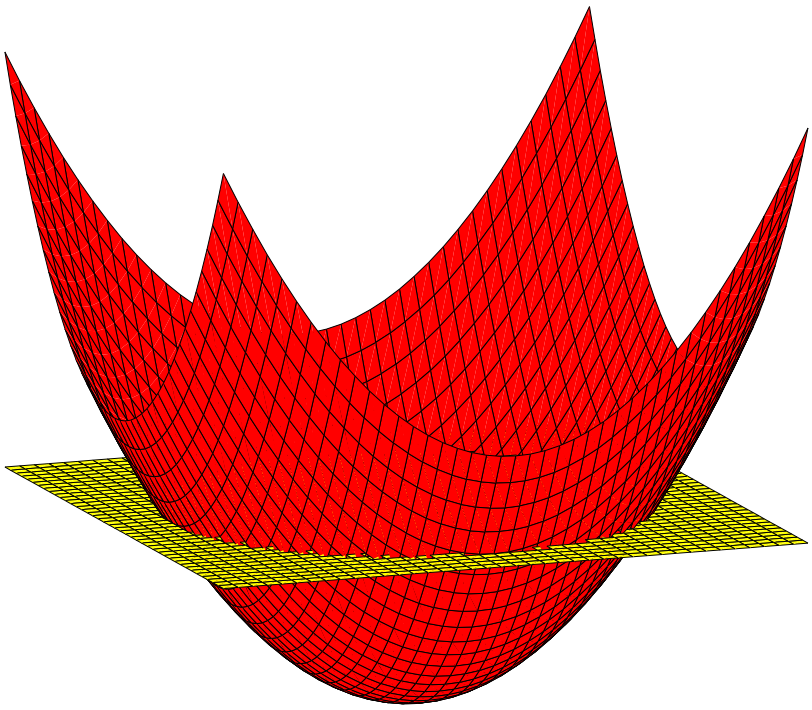

Figure 11: Singular values of the "unperturbed" matrix A.

clearly shows that it would be vain to attempt locating coalescing points in such cases.

3.2 Two parameters Case

In the next example we illustrate how a degenerate situation for a two parameter function breaks down under perturbation.

Example 3.2.1. Let $x_1, x_2 \in [-1, 1]$,

$$A(x_1, x_2) = \begin{bmatrix} x_1^2 + x_2^2 & 0 \\ 0 & .64 \end{bmatrix}.$$

The matrix A is real symmetric and positive definite, so its singular values coincide with its eigenvalues. The two singular values of A are given by a paraboloid-shaped surface and a horizontal flat plane, which coalesce along the circle of radius .8 centered at the origin, see Figure 11. This is clearly a degenerate situation since, by the result of von Neumann and Wigner, singular values of a 2-parameter dependent matrix

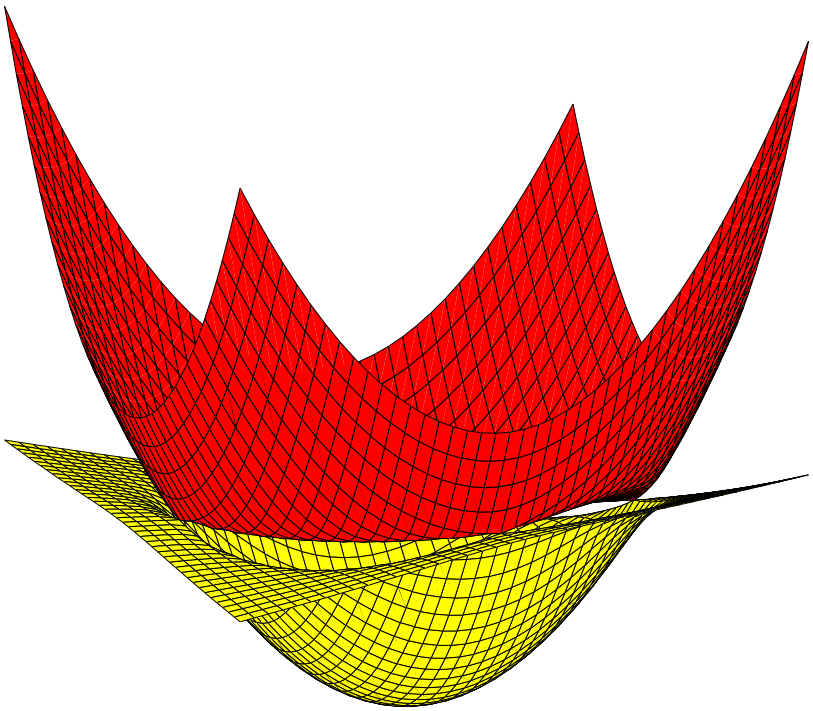

Figure 12: Singular values of the "perturbed" matrix $A + E$.

function are expected to coalesce at isolated points (see considerations at the end of Remark 2.1.17) . Just as in Example 3.1.1, let us perturb the matrix function A by adding a small perturbation matrix E to it. As Figure 12 shows, after the perturbation the singular values of $A + E$ do actually coalesce only at four isolated points (three of which are visible in the figure). Zooming-in on one of the coalescing points (see Figure 13) reveals that the two surfaces of singular values intersect in a double-cone fashion. This is another example of a conical intersection, of which we have reported in Section 1.3.1.

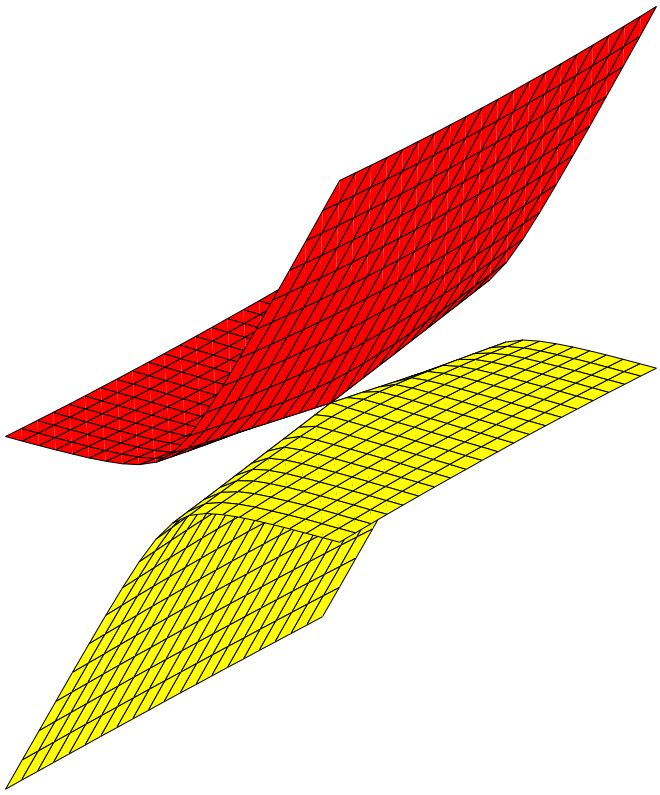

Figure 13: Zoom-in on a "conical intersection".

CHAPTER IV

ALGORITHMS FOR THE COMPUTATION OF COALESCING POINTS: TWO-PARAMETER CASE

In this Chapter, we describe and implement algorithms which aim at locating points in a simply connected domain Ω where the singular values of a two-parameter matrix function coalesce. We will restrict consideration to the case when Ω is a rectangle: $\Omega = \{x = (x_1, x_2) \in \mathbb{R}^2 \ : \ a \leq x_1 \leq b \ , \ c \leq x_2 \leq d \ \}$. The case of a triangular region is handled in a similar way, and the case in which Ω is a simply connected region with piecewise linear boundary is also handled with no major modifications. The case in which Ω is a disk is also simple, and one of our examples at the end considers this case. Through appropriate subdivisions, and reparametrizations, we can, in principle, tackle more complicated situations. We will be concerned only with the SVD of general matrices; the same algorithms developed here also apply to the eigendecomposition of symmetric matrices.

Our algorithm consists of two main modules: localization and zoom-in. Throughout this Chapter, we assume that coalescing of singular values occurs only at isolated points, as generically expected.

4.1 Localization

The localization module attempts to locate subregions of Ω where there is a *coalescing point*. By this, we mean a conical intersection, or more generally an intersection of odd multiplicity, of singular values. The technique we use is based on the theoretical results given in Section 2.2, see Theorem 2.2.13, which can be rephrased as follows: The presence of a coalescing point inside a region bounded by a closed curve triggers

period doubling of the orthogonal factors of the continuous SVD of A computed along the curve. This period doubling phenomenon is what we will use as evidence for the existence of a coalescing point inside a certain region.

4.1.1 Description of the Algorithm

Let us describe the localization module of the algorithm in more detail. First, we subdivide Ω in a cartesian grid of $N \times M$ boxes, not necessarily all of the same size. That is, we let $\Omega = \cup_{i=1:N,j=1:M} B_{ij}$, where $B_{ij} = \{(x_1, x_2) \ : \ x_1^{(i)} \leq x_1 \leq x_1^{(i+1)} \ , \ x_2^{(j)} \leq x_2 \leq x_2^{(j+1)} \}$, $i = 1, \ldots, N$, $j = 1, \ldots, M$. Let us call $x_{i,j} = (x_1^{(i)}, x_2^{(j)})$, and similarly $x_{i+1,j}$, $x_{i,j+1}$ and $x_{i+1,j+1}$, the vertices of the box B_{ij}.

Remark 4.1.1. Since coalescing points are isolated, we may as well assume that no coalescing point is on the grid itself. However, this eventuality (albeit not to be expected) would actually be easily detected; see Remark 4.1.2 below.

The idea is to "sweep" the grid, searching for boxes inside which there is a coalescing point. To decide upon the presence in a box of a coalescing point, we monitor whether or not the following condition is satisfied: Going around the box once, the orthogonal factors of a continuous SVD at the beginning of the loop and those at the end of the loop do not match. Here is how we do it.

Let us consider the general box B_{ij}. We think of the boundary of B_{ij} as made up of two L-shaped paths, Γ_1 joining $x_{i,j}, x_{i+1,j}, x_{i+1,j+1}$ and Γ_2 joining $x_{i,j}, x_{i,j+1}, x_{i+1,j+1}$; see Figure 14.

Note that both paths have the same starting and end points. We compute[1] the continuous SVD of A along both paths, starting from the same reference decomposition $A(x_{i,j}) = U_0 \Sigma_0 V_0^T$, and record the orthogonal factors U at the end of each path, call them $U_1(x_{i+1,j+1})$ and $U_2(x_{i+1,j+1})$. At this point, the two orthogonal factors are compared. If they match, nothing can be inferred about the existence of a coalescing

[1] see Section 4.1.2 for details about how this will be done.

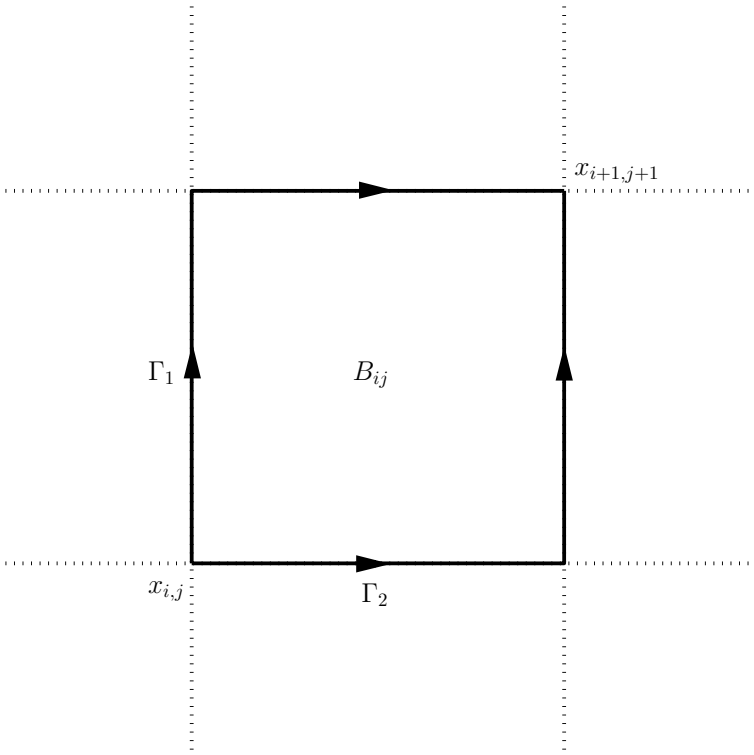

Figure 14: The paths Γ_1 and Γ_2 in a typical box.

point inside the box, and we either refine the box further or –if the size of the box is deemed sufficiently small– we declare that there is no coalescing point within the desired resolution. On the other hand, if they do not match, we look at the matrix $D := \mathrm{diag}\big(U_1(x_{i+1,j+1})^T\ U_2(x_{i+1,j+1})\big)$, which contains the information about what pair of singular values have coalesced inside the box under consideration; this is the same D as in Theorems 2.2.6 and 2.2.13.

This procedure is followed until all boxes have been explored. Information about which pairs of singular values have coalesced in which boxes is saved and will serve as input for the zoom-in module.

4.1.1.1 Local Refinement

Obviously, the localization procedure above can be implemented so to refine a box where there is a coalescing point. For example, progressively halving each vertical/horizontal side ("bisection" approach). By doing so, in principle, one would be able to locate a coalescing point with arbitrary degree of accuracy. Still, it is pretty obvious that this refinement process will converge only linearly, and this is one of the reasons why we prefer a quicker zoom-in technique once we have a good initial guess.

4.1.2 1-d Solver

A separate component of our techniques rests on the ability to compute the continuous SVD along the sides of a box, say the paths $\Gamma_{1,2}$ of which above. We describe the procedure relatively to computation of the path of SVDs along Γ_1 (obviously, the same applies to Γ_2). First of all, we force the path to step exactly (from x_{ij}) at both $x_{i+1,j}$ and to $x_{i+1,j+1}$, so that it is enough to describe the procedure on the line segments between these vertices, say between x_{ij} and $x_{i+1,j}$. We can thus think that we need to compute a smooth path of SVDs of the $\mathcal{C}^e([0,1],\mathbb{R}^{n\times n})$ function A, and we want $\mathcal{C}^e$ factors U, Σ, V. (The normalization to $[0,1]$ is for convenience only, obviously the physical interval in the case under consideration is the difference of the first components of $x_{i+1,j}$ and x_{ij}.)

It is absolutely crucial that the SVD be continued along each path in a continuous way, as otherwise we cannot make reliable inferences at the end point $x_{i+1,j+1}$. Assuming distinct singular values, it is easy to produce smoothly varying factors along the path of SVDs. Consider the following.

Let $0 = t_0 < t_1 < \cdots < t_N = 1$ be a partition of $[0,1]$, and let $h_j = t_{j+1} - t_j$, $j = 0, 1, \ldots, N-1$. In practice, the step-sizes h_j, $j = 1, \ldots, N-1$, will be found adaptively, see below. The first step-size h_0 is arbitrary. Our code takes $h_0 = 10^{-3}$ unless a different value is given in input. Anyway, the adaptive step-size strategy

quickly adjusts it. We have $A(0) = U_0\Sigma_0 V_0^T$ as our starting decomposition at t_0.

The basic technique consists of the following steps, which have been implemented by the second and third authors of [8][2].

1. Using canned software, we compute the SVD at the points t_j, $j = 1, 2, \ldots, N$, that is $A(t_j) = U_j\Sigma_j V_j^T$. Since the singular values of $A(t_j)$ are distinct, and ordered along the diagonal of Σ, then the orthogonal factors U_j and V_j are uniquely determined except for joint changes of signs of the columns of U_j and V_j. In order to guarantee continuity of the factors, we choose the signs of these columns in agreement with the signs of the columns of the factors in the *predicted* values at t_j, call these $U_j^{(0)}, V_j^{(0)}$. The latter values, the predicted values, are obtained by extrapolating at t_j the line through (t_{j-2}, U_{j-2}) and (t_{j-1}, U_{j-1}), and similarly for V. Then, we adjust the signs of U_j so that $(U_j e_k)^T (U_j^{(0)} e_k) > 0$, for all $k = 1, \ldots, n$. For the first step, that is at t_1, the predicted values are just the old values at t_0.

2. To choose the step-sizes adaptively, we use the same strategy of [8]. For simplicity, here below we omit the index j indicating the gridpoint t_j at which we are. We have predicted values $U^{(0)}, \Sigma^{(0)}, V^{(0)}$, and *corrected* values U, Σ, V, so that we can compute the errors $\|\Sigma^{(0)} - \Sigma\|$, $\|U^{(0)} - U\|$ and $\|V^{(0)} - V\|$. Next, we compute the weighted norm $\rho_\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n} (\frac{\sigma_i^{(0)} - \sigma_i}{\epsilon_r |\sigma_i| + \epsilon_a})^2}$, where ϵ_r and ϵ_a are relative and absolute error tolerances, and analogously we compute ρ_U, ρ_V. [Typically the values $\epsilon_r = \epsilon_a = 10^{-3}$ are sufficient, but sometimes we have to decrease this value, if required to do so by fast variation of the orthogonal factors.] Then, we set $\rho = \max(\rho_\sigma, \rho_U, \rho_V)$ and $h_{\text{new}} = \frac{h}{\sqrt{\rho}}$. If $\rho \leq 1.5$, the step is accepted, otherwise it is rejected. In all cases, h_{new} is used as the new steplength. Finally, the last step-size is chosen so to land exactly at $t_N = 1$.

Remark 4.1.2. If there are coalescing points along the step we take (on the line segment x_{ij} to $x_{i+1,j}$), they can be easily detected by the 1d-solver, by looking at the ordering of the singular values in the predicted $\Sigma_j^{(0)}$: If these are not ordered, then

[2]We thank M.G. Gasparo and A. Papini for letting us use a `Matlab` version of their program

we can infer which pair must have exchanged ordering and –as in [8]– we can trigger a one-dimensional secant search to locate the coalescing point precisely.

4.1.3 Implementation

We begin with a cartesian grid, which defines boxes B_{ij}, $i = 1, \ldots, N$, $j = 1, \ldots, M$. Our basic algorithm consists of going through the grid by covering one box at the time. In our implementation, we do this horizontally Left-to-Right and vertically Bottom-to-Top, across the grid. We transverse each edge of any box only once, and never compute twice the SVD at any given point.

As soon as all boxes are covered, and assuming that a coalescing value has been located inside a certain box, we either refine the box or trigger the zoom-in module. We found it effective to refine until the box is sufficiently small that Newton's method (on which the zoom-in module is based) can be estimated to converge quickly (say, in 4 to 5 iterations); see Example 4.3.2. Assuming that the boxes are squares, a good rule of thumb we have adopted would be to refine until the center of the box has distance from the vertices between 1/10 and 1/100. An important aspect to consider is that there is a trade-off between refining so to get closer to the coalescing point (and then zoom-in on it quickly) and expense. In fact, recall Remark 3.1.2, the computational work increases considerably near coalescing points, and so progressive refinements become increasingly more expensive because of the fast variation of the orthogonal factors near a coalescing point; see the considerations at end of Example 4.3.1.

Remark 4.1.3. As already remarked, the workhorse of our algorithms is the adaptive continuation code which computes the smooth SVD along the edges of the boxes. Some tweaks to this code have proved useful in order to make sure that it is particularly fit to our scope. For example, the localization method visits the boxes one at the time, which forces frequent restarting of the 1d-solver. Starting the continuation

along each edge with the same (most likely unnecessarily tight) initial step-size would result in an unpleasant overhead for the adaptive step-size selection. To avoid this, the 1d-solver has been implemented so to render, at the end of the continuation of, say, the edge from x_{ij} to $x_{i+1,j}$, the final step-size that would have been used if it did not have to step exactly at $x_{1+1,j}$. This step-size is then used as initial step-size for the continuation of the edge from $x_{i+1,j}$ to $x_{i+2,j}$. This strategy allows to save a substantial amount of unnecessary continuation steps.

4.2 Zoom In

Given a box B_{ij} in Ω and a pair of consecutive singular values which are known to coalesce in B_{ij}, we want to accurately locate the values of the parameter where the singular values coalesce.

Assume a coalescing between σ_k and σ_{k+1} has been detected in the box B_{ij}. Then we apply Newton's method to minimize the nonlinear functional

$$f(x) = (\sigma_k(x) - \sigma_{k+1}(x))^2 .$$

To be precise, since the function f is a smooth paraboloid-shaped function, we look for a solution of the problem $\nabla f(x) = 0$, rewritten as $F(x) = 0$, where $F : \mathbb{R}^2 \to \mathbb{R}^2$. On this systems, we perform Newton's method, using the center of the box as the initial guess. Obviously, the method will converge with quadratic rate if the initial guess is close enough to the exact solution.

4.2.1 Implementation

We summarize the algorithm below. As `INPUT`, we give a convergence tolerance `TOL` (we use `TOL` = 100 `EPS`, where `EPS` $\approx 2.2 \times 10^{-16}$ is the machine precision), and a maximum number of iterations `MAXIT` (we use 10), as well as the box B_{ij} under scrutiny and the pair of singular values which are known to coalesce in the box; say, $(k, k+1)$.

J=0, $x = \text{center}(B_{ij})$, $y = (1,1)$.

While $\min\{\|F(x)\|, \|y\|\} >$ TOL, and J<MAXIT, do

$DF(x)y \;=\; -F(x)$ [Solve for y]

$x = y + x$, J=J+1, enddo

Note that the correction vector y has been initialized to $(1,1)$ so to correctly enter the while cycle.

Remarks 4.2.1. Some observations are in order.

(i) First of all, since the singular values themselves are not expected to be differentiable, while $f(x)$ is, we form approximations of $F(x)$ and $DF(x)$ by differencing on f, as follows. For the gradient $F(x)$, we approximate it with the function $\widetilde{F}(x)$ obtained using 2-nd order divided differences:

$$F(x) \approx \widetilde{F}(x) = \begin{bmatrix} \big(f(x_1+h, x_2) - f(x_1-h, x_2)\big)/(2h) \\ \big(f(x_1, x_2+h) - f(x_1, x_2-h)\big)/(2h) \end{bmatrix} . \tag{8}$$

For the second derivatives making up $DF(x)$ we also use centered differences approximations. E.g.,

$$f_{x_1,x_1} \approx \big(f(x_1+h, x_2) - 2f(x_1, x_2) + f(x_1-h, x_2)\big)/h^2 ,$$

and

$$f_{x_1,x_2} \approx \frac{f(x_1+h, x_2+h) - f(x_1+h, x_2-h)}{4h^2} - \frac{f(x_1-h, x_2+h) + f(x_1-h, x_2-h)}{4h^2} .$$

It is a well known fact (easily proved using Taylor's expansion) that the above divided differences provide an approximation of order h^2 of the relevant derivative. But, besides the discretization error, we have to take into account also the *round-off* error, which, in the computation of the divided difference's formula,

is proportional to $\texttt{EPS}/h$. It is common practice in numerical analysis to choose a discretization step h that minimizes the sum of the two errors just considered. With this purpose in mind, we choose $h = \sqrt[3]{\texttt{EPS}}$ as discretization step in all the approximations above. The objective function $\widetilde{F}$ defined in Equation 8 will therefore be within $h^2 = \sqrt[3]{\texttt{EPS}^2} \approx 3.7 \times 10^{-11}$ of the real gradient F. We stress that $\widetilde{F}$ is the actual objective function that is brought to zero by the Algorithm described in Section 4.2.1. As a consequence, we will expect our zoom-in technique to converge to points where singular values coincide only up to a relative error of size 10^{-11}.

(ii) At any Newton step, it is necessary to evaluate the function f at 9 distinct points. Therefore, any Newton step requires (and costs as much as) nine SVDs. At present, we have not implemented any strategy to reduce this cost, since we have been chiefly interested in studying feasibility and reliability of the technique on small dimensional problems; of course, for large dimensional problems, one will need to implement new techniques whereby only the relevant singular values' information is computed.

(iii) Newton's method occasionally converges to a solution outside of the box B_{ij}, or fails to converge. This usually betrays a poor initial guess and triggers a refinement procedure for the box B_{ij}.

4.3 Examples

In this section we exemplify performance of the algorithms on several test problems. We highlight the following aspects: (i) Impact of the initial subdivision; (ii) Impact of refinement on the zoom-in module; (iii) Adaptation to a different geometry (non-rectangular); (iv) Performance of the algorithms on degenerate cases. All algorithms have been implemented in `Matlab` and run on a Laptop computer with clock speed of 2.2 GHz.

Example 4.3.1. Let $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ -1 \leq x_1 \leq 1 \ , \ -1 \leq x_2 \leq 1 \ \}$,

$$B(x_1, x_2) = \text{diag}(x_1^2 + x_2^2 \ , \ .81 \ , \ .36),$$

$$C = \begin{bmatrix} -0.179... & -0.294... & -0.722... \\ 0.787... & 0.626... & -0.594... \\ -0.884... & -0.980... & -0.602... \end{bmatrix},$$

$$E(x_1, x_2) = .5(x_1 + x_2)(x_1 + 1/3)\, C\,,$$

and $A(x_1, x_2) = B(x_1, x_2) + E(x_1, x_2)$, for all $(x_1, x_2) \in \Omega$. The entries of C have been chosen at random (uniformly distributed in $[-1, 1]$). Just as in Example 3.2.1, the matrix function A has been built as the sum of a matrix B, whose singular values coalesce in a degenerate way (along two circles centered at the origin of radius .6 and .9) and a perturbation matrix E. In order to force the existence of some coalescing points, the perturbation matrix E has been chosen so that $E = 0$ along two lines in the parameter space, which intersect the above-mentioned circles at eight points. Therefore, at least eight coalescing points are known to exist in Ω (their coordinates may even be easily computed by hand). Here we report on a few experiments using different choices of grids (all uniform). First, we run the localization module subdividing Ω into a 2×2 grid. Coalescing points for the pair (σ_2, σ_3) are detected in boxes $B_{1\,1}$ and $B_{2\,1}$, see Figure 15. Switching to a 10×10 grid, coalescing points are detected in 10 boxes, see Figure 16. [Note that 8 coalescing points had gone undetected through the 2×2 grid due to the fact that they occur in pair over each box.] We run the zoom-in module on each of the ten boxes, but we fail to converge to the coalescing points in $B_{7\,2}$ and $B_{4\,8}$. Finally, we repeat the procedure over a 100×100 grid. No further coalescing points are detected, but now the zoom-in module converges to all points. We illustrate the results obtained in Tables 1, 2, 3. Each line of the first three tables shows, in order: Box B_{ij} where a coalescing is detected, pair of singular values that coalesce, coordinates of the point where the zoom-in module has converged

(only the first four decimal digits are shown), number of Newton's iterates performed by the zoom-in module, actual difference $\Delta\sigma$ between the two singular values at the computed coalescing point.

A few observations follow in order:

(i) The average number of Newton's iterates for the case of the 10×10 grid is 7. With the 100×100 grid, where a better initial guess is provided (guaranteed to be within $\sqrt{2}/100 \approx 0.0141$ of the coalescing point), that value goes down to 5.3. These numbers betray the quadratic convergence rate of our zoom-in method.

(ii) The values of $\Delta\sigma$ are consistently in the order of 10^{-10}–10^{-11}. This fact has been predicted and confirms our analysis in Remark 4.2.1 (i).

We recall that our localization algorithm devolves an essential part of the work to a 1d-solver that computes the smooth SVD along the edges of the boxes. In Table 4, we present, in a comparison chart, workload and execution time required by each of the three tests considered in this example. Each line of Table 4 displays, in order: grid used in the test, total number of steps performed by the underlaying SVD continuation code, cpu-time required to perform the computation. We observe that the cpu-time increases linearly with the number of continuation steps, while it doesn't seem to increase as much with the number of boxes involved. This suggests that the most conspicuous share of work is indeed performed by the SVD continuation code.

We conclude this example by showing how the workload is distributed over the region Ω. We do this in the 100×100 grid case. For each box B_{ij}, we count the total number of steps performed by the SVD continuation code along the four edges of the box, and interpret it as the height corresponding to the center of the box. We represent these data in a 3D picture, displayed in Figure 17, where Ω lies on the xy plane. Several contour levels are drawn, each at the corresponding height. In this

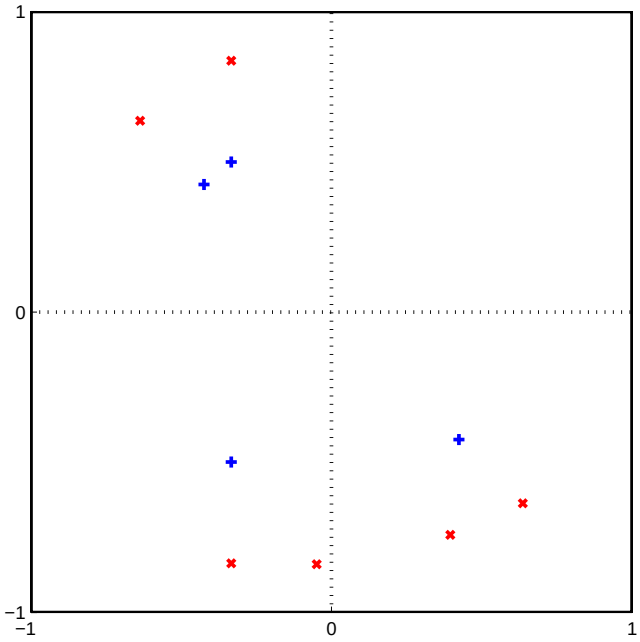

Figure 15: 2×2 grid; coalescing points are indicated with a $\times$ ($\sigma_1 = \sigma_2$) or a $+$ ($\sigma_2 = \sigma_3$)

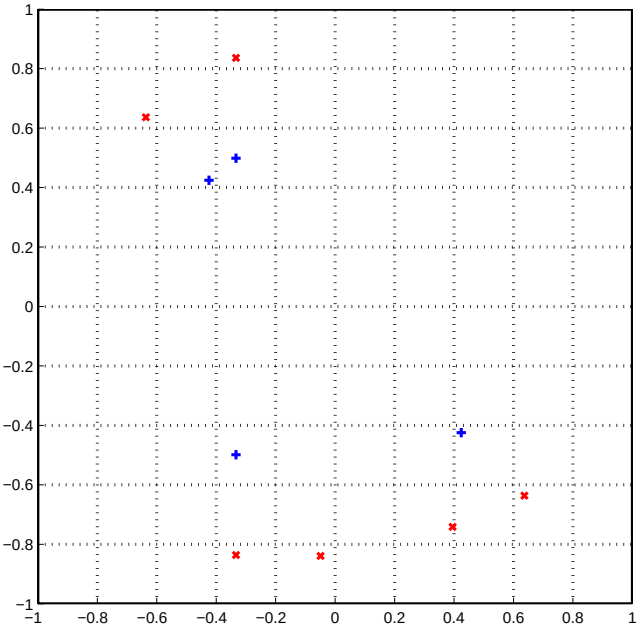

Figure 16: 10×10 grid; coalescing points are indicated as in Figure 15

Table 1: Performance on the 2×2 grid.

box	pair	point	iterates	$\Delta\sigma$
$B_{1\,1}$	(σ_2, σ_3)	$(-0.3333, -0.4989)$	7	5.6459×10^{-11}
$B_{2\,1}$	(σ_2, σ_3)	$(0.4243, -0.4243)$	6	5.9542×10^{-11}

Table 2: Performance on the 10×10 grid.

box	pair	point	iterates	$\Delta\sigma$
$B_{4\,1}$	(σ_1, σ_2)	$(-0.3333, -0.8360)$	8	4.9633×10^{-11}
$B_{5\,1}$	(σ_1, σ_2)	$(-0.0490, -0.8390)$	7	4.7905×10^{-11}
$B_{7\,2}$	(σ_1, σ_2)	–	–	–
$B_{9\,2}$	(σ_1, σ_2)	$(0.6364, -0.6364)$	7	3.2478×10^{-11}
$B_{4\,3}$	(σ_2, σ_3)	$(-0.3333, -0.4989)$	6	5.6459×10^{-11}
$B_{8\,3}$	(σ_2, σ_3)	$(0.4243, -0.4243)$	6	5.9542×10^{-11}
$B_{3\,8}$	(σ_2, σ_3)	$(-0.4243, 0.4243)$	7	7.1388×10^{-11}
$B_{4\,8}$	(σ_2, σ_3)	–	–	–
$B_{2\,9}$	(σ_1, σ_2)	$(-0.6364, 0.6364)$	7	1.1888×10^{-10}
$B_{4\,10}$	(σ_1, σ_2)	$(-0.3333, 0.8360)$	8	1.2190×10^{-10}

Table 3: Performance on the 100×100 grid.

box	pair	point	iterates	$\Delta\sigma$
$B_{34\,9}$	(σ_1, σ_2)	$(-0.3333, -0.8360)$	5	4.9632×10^{-11}
$B_{48\,9}$	(σ_1, σ_2)	$(-0.0490, -0.8390)$	5	4.7906×10^{-11}
$B_{70\,13}$	(σ_1, σ_2)	$(0.3951, -0.7417)$	5	2.8978×10^{-11}
$B_{82\,19}$	(σ_1, σ_2)	$(0.6364, -0.6364)$	5	3.2478×10^{-11}
$B_{34\,26}$	(σ_2, σ_3)	$(-0.3333, -0.4989)$	5	5.6459×10^{-11}
$B_{72\,29}$	(σ_2, σ_3)	$(0.4243, -0.4243)$	5	5.9541×10^{-11}
$B_{29\,72}$	(σ_2, σ_3)	$(-0.4243, 0.4243)$	5	7.1388×10^{-11}
$B_{34\,75}$	(σ_2, σ_3)	$(-0.3333, 0.4989)$	6	3.7166×10^{-11}
$B_{19\,82}$	(σ_1, σ_2)	$(-0.6364, 0.6364)$	6	1.1888×10^{-10}
$B_{34\,92}$	(σ_1, σ_2)	$(-0.3333, 0.8360)$	6	1.2190×10^{-10}

Table 4: Comparison of workload and execution time.

grid	continuation steps	cpu-time
2×2	1176	0.78 secs
10×10	6170	2.83 secs
100×100	71304	40.64 secs

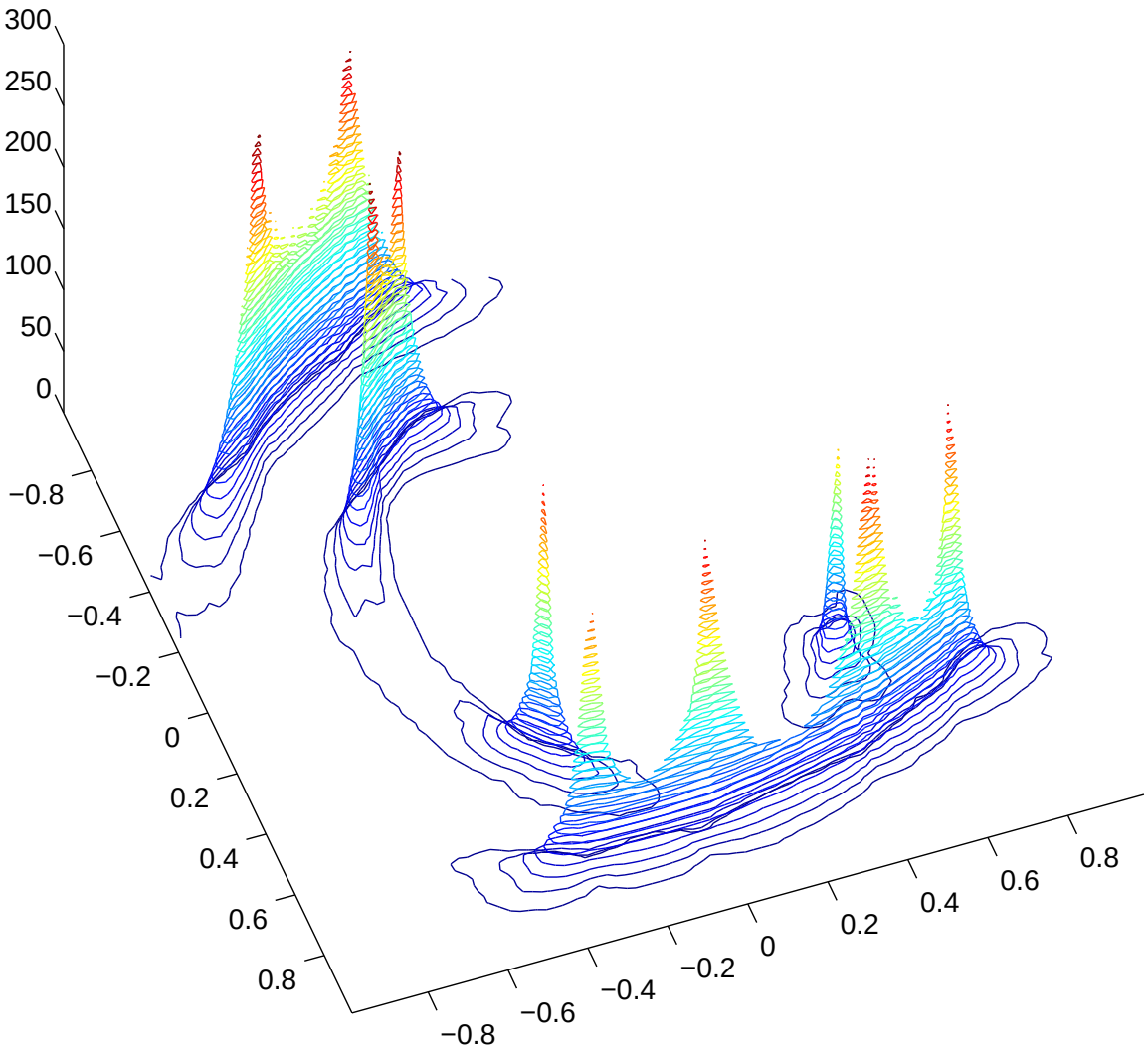

Figure 17: Distribution of the workload over the region Ω.

case, the minimum number of steps/box is 8, while the maximum is 288. Ten peaks are clearly noticeable in the picture: They all correspond to coalescing points (see Figure 16). Most of the computational expense is concentrated around coalescing points, as ought to be expected by virtue of Remark 3.1.2. With this in mind, the strategy we have adopted in our experiments has been to guess a grid-size that is just enough for the zoom-in iterates to converge. Refining any further would add unnecessary weight to the algorithm.

Example 4.3.2. Let us consider again the function A as in Example 4.3.1. In that Example, the 10×10 grid was enough to detect all coalescing points but not to zoom-in on all of them. Now, instead of switching to a finer grid over the entire region, we will perform a local refinement, only where needed. In this example, we do a bisection-type approach on box $B_{7\,2}$. We perform 5 successive bisections of each side of the box, thus getting within ≈ 0.0044 of the coalescing point (see shaded region on Figure 18 and beware of the coordinates values). Now we can safely run the zoom-in module and expect to converge to the coalescing point in at most 4 Newton iterates, as indeed happens.

Example 4.3.3. As already mentioned, our localization technique does not place, in principle, any restriction on the choice of the planar grid. Specific geometric properties of the domain of interest may suggest different choices of grids, other than cartesian. In this example, we consider the same matrix function A as in Examples 4.3.1 and 4.3.2, defined on the circular domain $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ x_1^2 + x_2^2 \leq 1 \ \}$. Given the particular shape of Ω, we decided to use the grid shown in Figure 19. Even though it is not cartesian, the grid can be assimilated to the union of (four) cartesian ones via a change to polar coordinates. Dimension and number of the "tiles" have been chosen so that the distance between the centroid of each tile and any point on it is not larger than 0.2. With this choice of grid, two coalescing points for the pair (σ_2, σ_3)

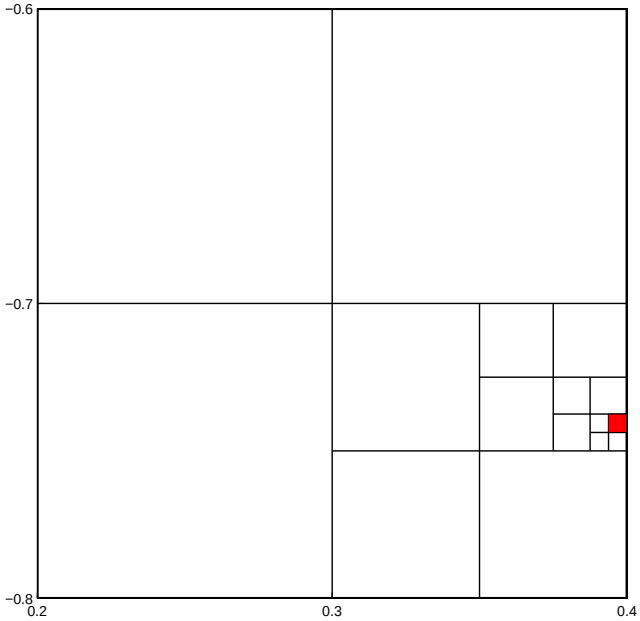

Figure 18: Local refinement on box $B_{7\,2}$

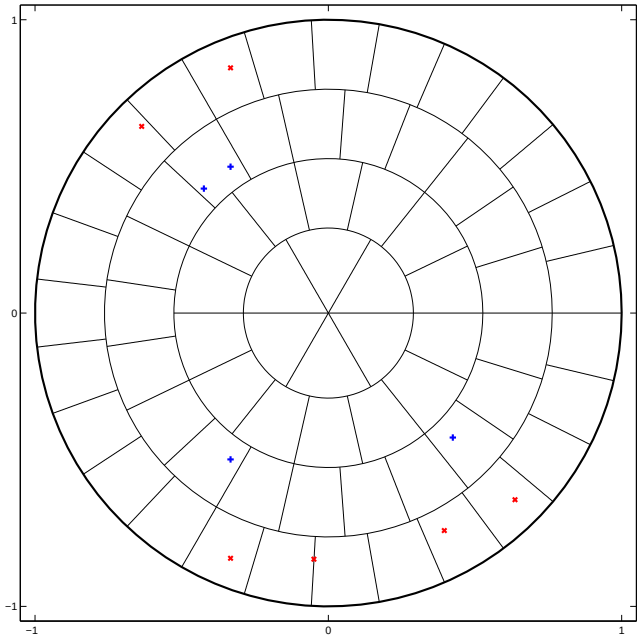

Figure 19: Non-cartesian grid on a circular domain; coalescing points are indicated as in Figure 15

go undetected. The localization module locates the remaining eight points, but the zoom-in module fails to converge to two of them.

Example 4.3.4. This is the first of two examples chosen to highlight what happens for degenerate (i.e. of multiplicity higher than two) intersections.

Let $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ -1 \leq x_1 \leq 1 \ , \ -1 \leq x_2 \leq 1 \ \}$ and

$$A_p(x_1, x_2) = \begin{bmatrix} x_2 + 2 & x_2 \\ x_2 & x_1^p + 2 \end{bmatrix},$$

where $p \geq 1$ is an integer. For all p, A_p is symmetric positive definite on Ω, so that its singular values coincide with its eigenvalues. Direct computation shows that its eigenvalues coalesce if and only if

$$\begin{cases} x_2 = x_1^p \\ x_2 = 0 \end{cases}, \tag{9}$$

i.e. if and only if $x_1 = x_2 = 0$. Following Definition 2.2.9, the multiplicity of this coalescing point is p if the order of contact of the two planar curves defined by the equations (9) is $p - 1$. Hence, in this example, the origin is a coalescing point of multiplicity p for A_p, for all $p \geq 1$. According to Theorem 2.2.13, a coalescing point does not produce any period doubling if its multiplicity is even. This causes our localization algorithm to fail to detect coalescing points of even multiplicity. We have tested our localization module on A_p for $p = 1, 2, 3$, and observed what was predicted by the Theorem: The coalescing point is correctly detected for $p = 1, 3$ while goes undetected for $p = 2$. We stress that, in general, one should not expect to have coalescing points of multiplicity $p \geq 2$ for matrix functions depending on two parameters, as the codimension of such phenomena is higher than two.

Example 4.3.5. Let $\Omega = \{(x_1, x_2) \in \mathbb{R}^2 \ : \ -1 \leq x_1 \leq 1 \ , \ -1 \leq x_2 \leq 1 \ \}$ and

$$A(x_1, x_2) = \begin{bmatrix} x_1 & x_1 + x_2 \\ x_1 - x_2 & x_2 \end{bmatrix}.$$

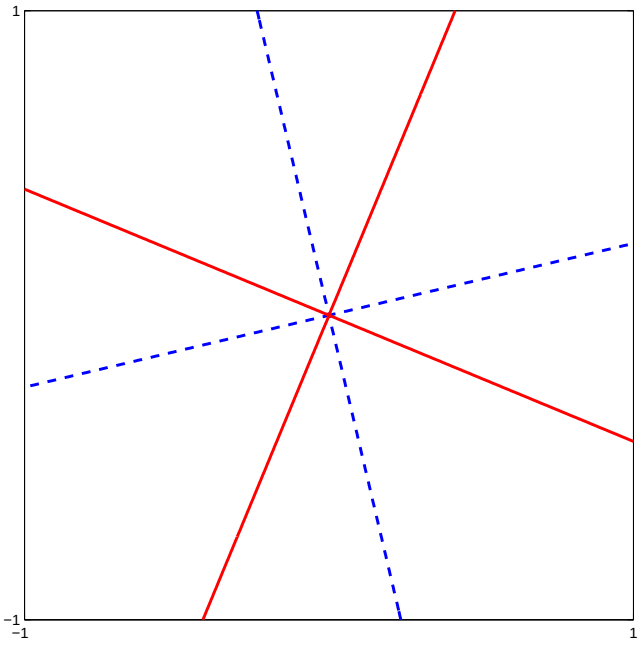

Figure 20: A codimension-4 coalescing: $\sigma_1 = \sigma_2 = 0$

The only coalescing point in Ω is the origin: $(x_1, x_2) = (0, 0)$, where we have $\sigma_1 = \sigma_2 = 0$. Therefore, we are in the degenerate situation considered in Remark 2.1.18. Our theoretical results in Chapter 2 do not cover this case, and we are interested in exploring this situation through our numerical techniques. We have launched our localization module, using only one box $B_{11} = \Omega$. No period doubling of the singular vectors is observed along the closed circuit. Therefore, the coalescing point goes undetected by our localization algorithm. The reason why this happens can be understood by looking at Figure 20. The blue (dashed) line represents the curve $a^2(x) + c^2(x) - b^2(x) - d^2(x) = 0$, the red (solid) line represents the curve $a(x)b(x) + c(x)d(x) = 0$. Considering the smooth SVD of A along the boundary of B_{11}, we notice that both the singular vectors cross each of the axis twice (recall the considerations in Example 2.2.15, where a symmetric eigenproblem is considered). Therefore, they must come back to their initial position.

CHAPTER V

EXTENSIONS: SEVERAL PARAMETERS

We have already pointed out that coalescing of singular values of real valued general matrix functions is a codimension-2 phenomenon. In the case of A depending on several, $k \geq 3$, parameters, this implies that singular values are expected to coalesce along $k-2$ dimensional manifolds. In particular, we expect the coalescing to occur along curves if A depends on three parameters, surfaces if A depends on four parameters, etc.

In this chapter, we study the case of matrix functions depending on three parameters. We start with real valued functions, tackling the problem of how to numerically compute curves along which singular values coalesce. Then, after having illustrated the main differences between real and complex cases, we show how the numerical algorithms introduced for the real case can be used to handle the problem of detecting/approximating coalescing points in the complex case. To witness, we present two examples in which we exemplify the numerical procedures that have been introduced. We discuss only the case of the SVD, the case of the eigenproblem being essentially identical.

Let us consider a $\mathcal{C}^e$, $e \geq 1$, matrix function $A : \mathbb{R}^3 \to \mathbb{R}^n$. For such a matrix function, several "solution curves" are expected to exist for each one of the problems:

$$\sigma_1(x) = \sigma_2(x),\ \sigma_2(x) = \sigma_3(x), \ldots,\ \sigma_{n-1}(x) = \sigma_n(x)\,,$$

where $x = (x_1, x_2, x_3)$. Such curves are expected to be smooth, as they are –locally, and away from degenerate points– implicitly defined as zero set of smooth functions: $F(x) = 0$, where F is defined as in Theorem 2.1.14.

Each curve can be made up of several branches. Here we propose different numerical techniques aimed at detecting and computing those branches.

5.1 Continuation of a Curve of Coalescing Singular Values

The first technique we present is a numerical method for the computation of a branch of the curve along which two singular values of a matrix A coalesce:

$$\{x \in \mathbb{R}^3 : \sigma_m(x) = \sigma_{m+1}(x)\}\,,$$

with $1 \leq m \leq n-1$. The method we propose resembles Keller's pseudo-arclength method for the continuation of implicitly defined curves, see [19, 14]. For this reason, we start with a brief overview of Keller's method.

5.1.1 Keller's Pseudo-Arclength Continuation

Let $G : \mathbb{R}^{n+1} \to \mathbb{R}^n$ be a $\mathcal{C}^1$ map and let 0 be a regular value for G. The zero set of G is a 1-dimensional differentiable manifold, whose maximal connected components are called branches; they are necessarily diffeomorphic with real intervals. Here the problem under consideration is that of computing solution branches of the equation

$$G(x) = 0\,.$$

This is a problem that arises naturally in the study of bifurcation of dynamical systems depending on a parameter:

$$\dot{x} = G(x, \lambda)\ (x \in \mathbb{R}^n, \lambda \in \mathbb{R}\,)\,,$$

therefore it is often presented in the form

$$G(x, \lambda) = 0\,, \tag{10}$$

where $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$ is considered as a distinguished parameter. By numerical continuation we mean a technique to find consecutive points on a solution branch of the previous equation.

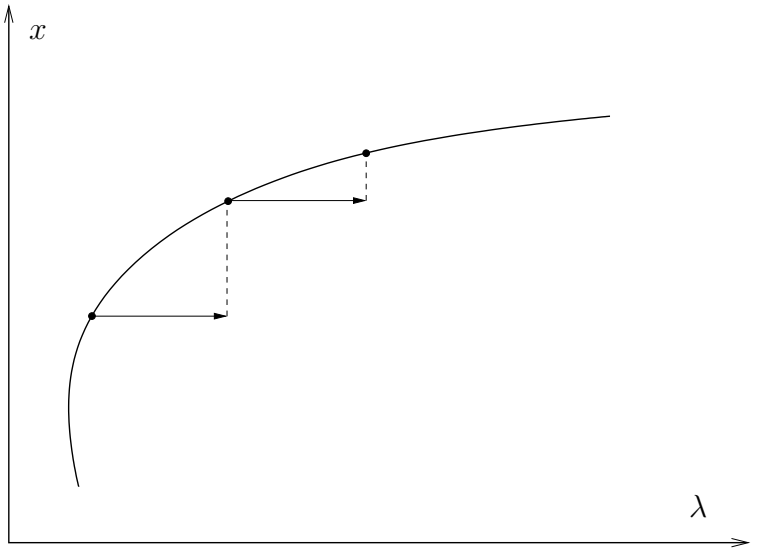

Figure 21: Continuation by natural parametrization (trivial predictor).

Suppose we already know a solution (x_0, λ_0), and want to find a new one, call it (x_1, λ_1). Several strategies have been proposed to perform this task. The most commonly adopted, and most successful, are those that go by the name of predictor-corrector techniques. Let us review a few. One possibility, the simplest, is to fix λ_1 close to λ_0, and then determine x_1 by solving

$$G(x, \lambda_1) = 0\,.$$

The equation is usually solved by Newton's method, using x_0 as initial guess. Geometrically, this amounts to approximating the curve first by a straight line (predictor step) and then correcting along the hyperplane $\lambda = \lambda_1$ (corrector step), see Figure 21. For its simplicity, the predictor just described is known as trivial predictor. Other choices of the predictor are possible; for instance, one could approximate the curve along the tangent line (see Figure 22), obtaining a better prediction of the next point on the branch

$$\widehat{x} = x_0 + \frac{\delta x}{\delta \lambda}(\lambda_1 - \lambda_0)\,,$$

to be more easily corrected via Newton's method.

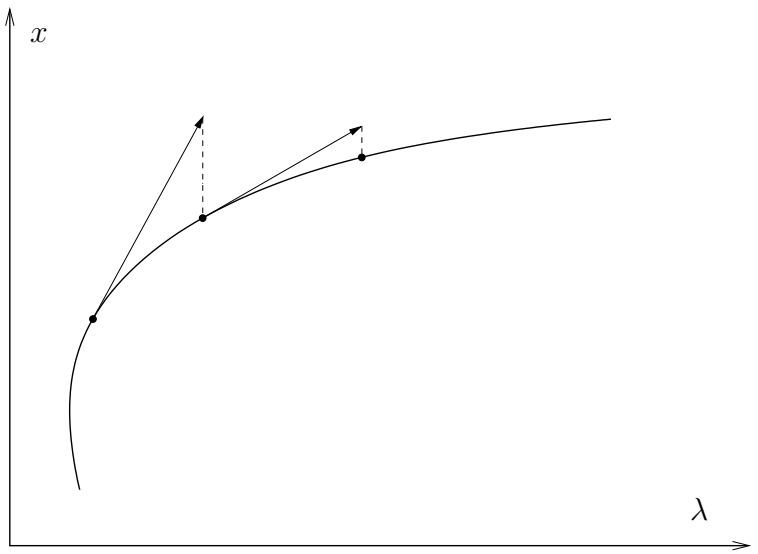

Figure 22: Continuation by natural parametrization (predictor along the tangent).

Left apart the choice of the predictor, the methods described above share a weakness: They suppose that the parameter λ can be used to parametrize the branch:

$$G(x(\lambda), \lambda) = 0\,.$$

For this reason, they perform the so called *continuation by natural parametrization*. Difficulties are encountered if the branch contains *fold points*, i.e. points (x, λ) along the branch where G_x is singular. In fact, λ is not a good parametrization in a neighborhood of such points, as the assumptions needed by the implicit function theorem to parametrize the curve in terms of λ fail to be satisfied. This is why a new method has been proposed, that uses the parametrization by arclength:

$$G(x(s), \lambda(s)) = 0\,.$$

With this choice of parametrization, a distinction between one variable (the parameter λ) and the others is no longer needed; so, we switch to the notation

$$G(x(s)) = 0\,,$$

where $x(s) \in \mathbb{R}^{n+1}$ for all s.

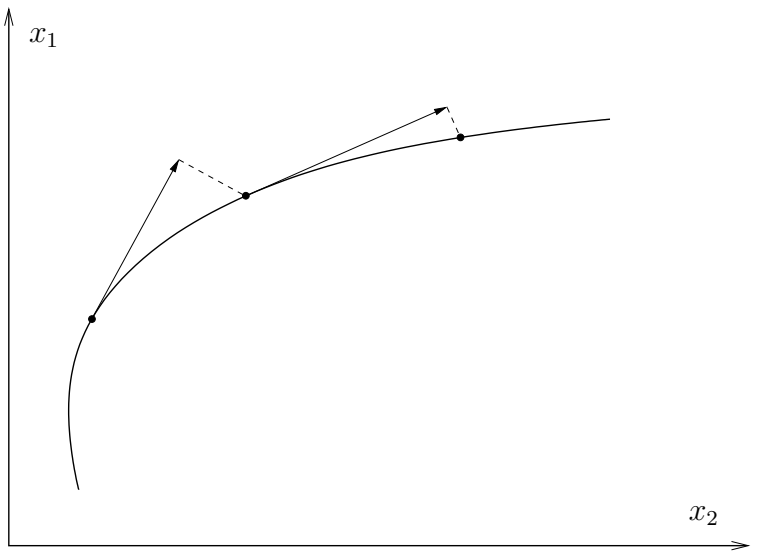

Figure 23: Keller's pseudo-arclength method.

Suppose that a point x_0, with unit tangent vector $\mathbf{t}_0$ is known, and suppose that a step-length Δs has been chosen. Then the predictor is defined as

$$\widehat{x} = x_0 + \Delta s\, \mathbf{t}_0\,.$$

This is called pseudo-arclength predictor because Δs approximates (along the tangent line) the arclength along the branch. The new solution to (10) is then sought on the hyperplane through $\widehat{x}$, orthogonal to $\mathbf{t_0}$; i.e. the next point x_1 is the solution of the following system of equations:

$$\begin{cases} G(x) = 0 \\ \mathbf{t}_0^T(x - \widehat{x}) = 0 \end{cases}$$

We have just described what is commonly referred to as Keller's method, see Figure 23.

5.1.2 Description of the Algorithm

After this digression on Keller's method, we are ready to describe our method. It is essentially an adaptation of Keller's method to our context. The geometric idea is

the same: prediction step along an approximate tangent (based on quadratic interpolation), correction step on the plane through the predictor, perpendicular to the tangent. More precisely, the prediction is corrected by means of our zoom-in module of Section 4.2.

Let $P(s)$, $s \in \mathbb{R}$, be a parametrization by arclength of the branch under consideration. Although we understand that a curve may be made up of different branches, we will use indifferently the words curve and branch, whenever this doesn't cause any confusion.

We will construct an approximation of the branch that consists of several successive points computed along the branch. Suppose three points on the branch are known: $P^{(k-2)}$, $P^{(k-1)}$ and $P^{(k)}$. Here we describe how we compute the next point $P^{(k+1)}$.

1. First, we construct a suitable quadratic interpolant φ, $\varphi(s) = a_2 s^2 + a_1 s + a_0$, to the three given points. Ideally, we want s to represent the arclength. Therefore, the interpolant φ is defined by the conditions:

$$\varphi\Big(- \big\|P^{(k)} - P^{(k-1)}\big\| - \big\|P^{(k-1)} - P^{(k-2)}\big\| \Big) = P^{(k-2)}\,,$$

$$\varphi\Big(- \big\|P^{(k)} - P^{(k-1)}\big\| \Big) = P^{(k-1)}\,,$$

$$\varphi(0) = P^{(k)}\,.$$

The interpolant is used in order to construct an approximation $\mathbf{n}^{(k)}$ of the tangent to the branch at the point $P^{(k)}$, i.e. we define

$$\mathbf{n}^{(k)} = \frac{\varphi'(0)}{\|\varphi'(0)\|}\,.$$

Let h be the step-size (its choice will be discussed later). Then, we define the predictor as

$$\widehat{P}^{(k+1)} = P^{(k)} + h\,\mathbf{n}^{(k)}\,.$$

Since the prediction is done along an approximation of the tangent line, the error of the predictor is of order h^2, if h is small. This can be easily verified from the Taylor expansion of $P(s)$.

2. Now consider the plane π through $\widehat{P}^{(k+1)}$, having $\mathbf{n}^{(k)}$ as its (unit) normal vector. The plane will intersect the curve if the step-size h and the curvature are not too large, see Figure 24. We think of the points on the plane as parametrized by:

$$\pi(t_1, t_2) = \widehat{P}^{(k+1)} + t_1\,\mathbf{q}_1 + t_2\,\mathbf{q}_2\,,$$

where $[\mathbf{q}_1,\ \mathbf{q}_2]$ is an orthonormal basis for π. There is no need for smoothness at this stage, and the dimension of the vectors involved is only three, therefore the orthogonal basis is computed simply using the `svd` routine from `Matlab`. Finally, the prediction $\widehat{P}^{(k+1)}$ is corrected by launching the zoom-in module of Section 4.2 for the function $A(\pi(t_1, t_2))$ (restriction of A to the plane π), using $(t_1, t_2) = (0, 0)$ as initial guess for the Newton's method. If the zoom-in module fails to converge, we halve the step-size h and go back to the prediction step. If Newton's method does converge, we declare it has converged to $P^{(k+1)}$. The convergence tolerance is provided by the user, see Section 4.1.3 for details.

3. The choice of the step-size h is very important: if it's too small, it will lead to a lot of unnecessary work; if it's too large, the correction algorithm may fail to converge, or converge to a different branch. We choose the step-size adaptively. Let $h_{\,\mathrm{old}}$ be the step-size used to compute $P^{(k)}$. The distance between predicted value at the k-th step, $\widehat{P}^{(k)}$, and corrected value, $P^{(k)}$, is used to update the step-size for the next step, $h_{\,\mathrm{new}}$, through the following formula:

$$h_{\,\mathrm{new}} = \tau\, h_{\,\mathrm{old}};\quad \tau = \sqrt{\frac{\mathtt{tol}}{\left\| P^{(k)} - \widehat{P}^{(k)} \right\|}}\,.$$

The formula above comes from the following considerations (and see Section 4.1.2). Let $s_k \in \mathbb{R}$ and $s_{k+1} = s_k + h$, with $h > 0$ small. From Taylor's expansion, we have:

$$P(s_{k+1}) = P(s_k) + h\,\dot{P}(s_k) + \frac{h^2}{2}\,\ddot{P}(s_k) + \mathcal{O}(h^3)\,.$$

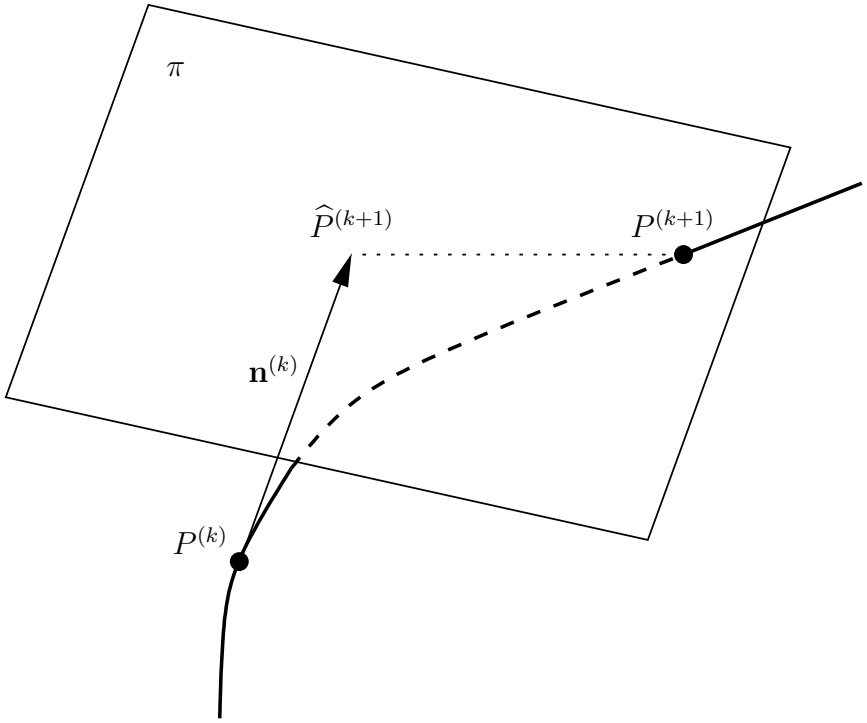

Figure 24: Continuation of a curve of coalescing points.

Now let $P^{(k)} = P(s_k)$ and $P^{(k+1)} = P(s_{k+1})$. In the equation above, $\dot{P}(s_k)$ can be replaced by $\left\|\dot{P}(s_k)\right\| \mathbf{n}^{(k)}$, and $\left\|\dot{P}(s_k)\right\|$ factored in the step h, so to obtain:

$$P^{(k+1)} = P^{(k)} + h\,\mathbf{n}^{(k)} + h^2 \frac{\ddot{P}(s_k)}{2\left\|\dot{P}(s_k)\right\|^2} + \mathcal{O}(h^3)\,,$$

from which, using $\widehat{P}^{(k+1)} = P^{(k)} + h\,\mathbf{n}^{(k)}$, and taking norms on both sides, we have:

$$\left\|P^{(k+1)} - \widehat{P}^{(k+1)}\right\| \approx h^2 \frac{\left\|\ddot{P}(s_k)\right\|}{2\left\|\dot{P}(s_k)\right\|^2}\,.$$

Finally, imposing

$$h_{\text{old}}^2 \frac{\left\|\ddot{P}(s_k)\right\|}{2\left\|\dot{P}(s_k)\right\|^2} = \texttt{tol}\,,$$

and letting $h_{\,\text{new}} = \tau h_{\,\text{old}}$, we obtain the given formula for τ. As a result of this formula, the step-size is shortened/streched, attempting to obtain an absolute error in the prediction that is close to `tol`.

If the step-size falls below a certain minimal threshold, the continuation is interrupted. Analogously, the step-size is not allowed to take values above a certain maximal threshold, to guarantee that the points $P^{(k)}$ produce a sufficiently faithful representation of the branch.

5.1.2.1 How to Start the Procedure

What has been just described is how the method works at regime. In fact, due to the quadratic interpolant adopted to construct the predictor, three points on the curve have to be known in order to start the procedure. Therefore, $P^{(0)}$, $P^{(1)}$ and $P^{(2)}$ must be handled in a different way. The first point on the branch, $P^{(0)}$, is assumed to be given. Then, the "trivial predictor" is used to construct $P^{(1)}$:

$$\widehat{P}^{(1)} = P^{(0)}\,.$$

Once $P^{(1)}$ has been computed, we can use a secant predictor

$$\mathbf{n}^{(1)} = \frac{P^{(1)} - P^{(0)}}{\|P^{(1)} - P^{(0)}\|}$$

$$\widehat{P}^{(2)} = P^{(1)} + h\,\mathbf{n}^{(1)}$$

and correct it to compute $P^{(2)}$. Now that three points are known, the quadratic corrector can be implemented.

One way to construct $P^{(0)}$ is to apply the localization and zoom-in modules studied in Chapter 4 to the restriction of the matrix function A to a suitable hyperplane. A simple but effective possibility is to keep one of the three variables frozen to a suitable value, and let the other two vary in a rectangle R. That said, to compute the entire curve, we do not want, and do not, assume that two of the variables are parametrized in terms of the other, as that may fail at fold points. Indeed, just as Kellers method, our method allows to continue a branch past a fold.

Remark 5.1.1. Of course, the quadratic interpolant is not the only possibility of

defining $\mathbf{n}^{(k)}$, the approximation of the tangent to the curve at $P^{(k)}$. We have experimented also with the use of the linear interpolator (secant) through all the steps, instead of just for the computation of $P^{(2)}$. At the end, we have opted for the quadratic one as it provides a better approximation.

5.1.3 Detection of a Different Curve

Suppose we are continuing a branch of the curve $\{x \in \mathbb{R}^3 : \sigma_m(x) = \sigma_{m+1}(x)\}$. An interesting problem is that of detecting other branches that pass close by this branch, relative to different pairs of singular values or possibly even to the same pair. Here we describe the method we have adopted to solve this problem. This can be considered as a separate module to be added to the second step in the continuation procedure described in Section 5.1.

After having completed step 2 of the continuation procedure, consider the square $S_\alpha = \{\pi(t_1, t_2) : (t_1, t_2) \in [-\alpha, \alpha] \times [-\alpha, \alpha]\}$, $\alpha > 0$, on the plane π. This square is centered on the "prediction" $\widehat{P}^{(k+1)}$. Now we simply search S_α, applying the localization module of Section 4.1 to the matrix function $A(\pi(\cdot, \cdot))$. If new coalescing points are detected, other than those for the pair (σ_m, σ_{m+1}), the zoom-in module is launched in order to approximate those points, and the relevant information is saved for future use. It is possible that more then one coalescing point is detected for the pair (σ_m, σ_{m+1}). In this case, after having run the zoom-in module, a vicinity check is performed in order to define the point $P^{(k+1)}$; the other points are marked as they could be part of a different branch or of the same one. The latter case can happen in correspondence of sharp turns. Through the parameter α, one can tune the size of S_α, so to choose whether to only detect branches that are close to the one under consideration, or search for ones that may be far apart from it. In principle, S_α can be searched through the use of an arbitrary grid. Nonetheless, we have chosen to avoid any subdivision (i.e., we use only one search box $B_{11} = S_\alpha$), as any further

refinement results in a dramatic drop in performance of the algorithm.

Remark 5.1.2. The possibility of implementing a "detection" feature is a peculiarity of our approach. In fact, our localization module allows us to literally swipe a given part of the plane perpendicular to the approximate tangent, in search for other coalescing points (each corresponding to the intersection with a possibly new curve); this was not possible in Keller's method as there is no natural way of detecting different branches in the case of curves implicitly defined by nonlinear equations. Therein, the equation of the plane perpendicular to the tangent can only be added to the equation $G(x, \lambda) = 0$.

5.1.4 Detection of Closed Loops

In general, a smooth curve $\{x \in \mathbb{R}^3 : \sigma_m(x) = \sigma_{m+1}(x)\}$ will be homeomorphic to a line or to a circle S^1. It is important to implement an algorithm aimed at distinguishing between the two cases, so to detect when the curve being continued is closed and to avoid tracing the same path over and over. This problem had already been faced, and solved, by Melville and Mackey in [22], in the context is the computation of two-dimensional manifolds. Our method is an adaptation of the "piercing computation" method therein. We now explain how we detect closed loops.

The one-dimensional continuation module described in 5.1 constructs points $P^{(0)}$, $P^{(1)}$, $P^{(2)}, \ldots$ along the curve. Eventually, for a closed curve, some pair $P^{(k)}$, $P^{(k+1)}$ may "bracket" $P^{(0)}$ in the sense of increasing arc-length, signaling that the branch being continued is a closed curve. We want to detect this event.

First, we construct an approximate unit tangent $\mathbf{n}_0$ to the curve at $P^{(0)}$. Then, we construct the plane π_0 through $P^{(0)}$ with $\mathbf{n}_0$ as normal vector, see Figure 25. Whenever a successive pair of points $P^{(k)}$, $P^{(k+1)}$ fall on two opposite sides of π_0, we construct the segment from $P^{(k)}$ to $P^{(k+1)}$ and compute the intersection of this segment with π_0. Finally, we "correct" this intersection, using it as initial value for

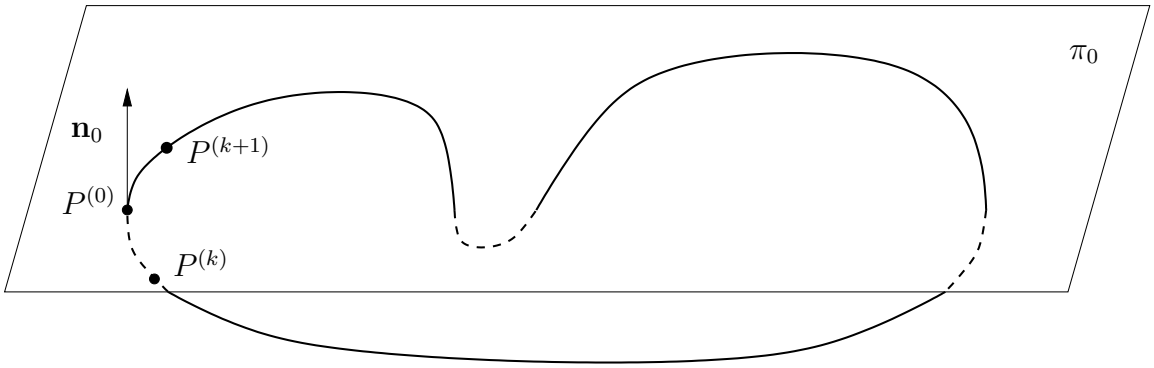

Figure 25: Piercing Computation.

the zoom in module of 4.2, applied to the function A restricted to the plane π_0. We declare the curve *closed* if this corrected piercing point is "very closed" to $P^{(0)}$, i.e. within a given tolerance based on the machine precision.

Remark 5.1.3. A complicated curve could wind back and forth across π_0 several times, but there is a natural way to rule out half of the crossings by an obvious parity check. Since, if the curve is closed, it must cross π_0 an even number of times, all crossings are enumerated and only the even-numbered are tested to check whether one is the closing point of the branch.

5.2 Extension: Complex Case

So far, we have been concerned only with the case of real matrix functions that depend on two, and more, parameters. In this section, we will consider the case of complex valued matrix functions. We will soon appreciate how this case is profoundly different, and needs a completely different approach.

The non-crossing rule by J. Von Neumann and E. Wigner states that, for a complex hermitian matrix, having a pair of coalescing eigenvalues is a codimension-3 phenomenon, and the same is true if we look at a coalescing pair of singular values for a complex, general (in particular, non symmetric) matrix. As one can easily anticipate, this jump in the codimension of the problem, from 2 to 3, has a strong impact on the scenario: Singular values are now expected to coalesce at isolated points for

3-parameters matrix functions, along curves for 4-parameters matrix functions, etc. On the other hand, coalescing of singular values for matrix functions that depend on 2 parameters is, now, non-generic and should not be expected. In this new setting, the problem that makes sense to study is that of how to detect, and approximate, coalescing points for the singular values of complex matrix functions depending on *three* parameters. The main difference, with respect to the case of real valued matrix functions, is that we are no longer looking at points on the plane, but rather at points in space. Therefore, the main idea that has pervaded our work throughout Chapters 2-4, i.e. that of enclosing the coalescing points with closed curves, has to be dropped, as it has no meaning in $\mathbb{R}^3$.

Even though we cannot rely anymore on the nice geometrical properties of curves on the plane, we still want to, and will, make use of the results we have derived so far for the case of real valued matrices. The example we are about to present is our first approach to the problem of detecting/approximating coalescing points for the singular values of complex matrix functions that depend on three (real) parameters. As we are about to see, the topological flavor of the techniques studied in Chapter 4 will be replaced by a completely new argument, that is perturbative in nature.

In the following, we give a description of our idea for the complex case, followed by an example of its application. Let us consider a differentiable matrix function $A : \mathbb{R}^3 \to \mathbb{C}^{n\times n}$. Our purpose is to transform the problem into a new one, that can be handled with the techniques in hand.

5.2.1 Description of the Method

1. The first thing we do is "double the problem", turning it into a real valued problem of twice the size. Let $A(x) = B(x) + iC(x)$, with $B(x), C(x) \in \mathbb{R}^{n\times n}$, for all $x \in \mathbb{R}^3$. Now let

$$M(x) = \begin{bmatrix} B(x) & -C(x) \\ C(x) & B(x) \end{bmatrix},$$

for all $x \in \mathbb{R}^3$. The following fact can be verified by direct computations:

$$(u_1^T + iu_2^T)\, AA^T (u_1 + iu_2) = \sigma^2$$

if and only if

$$\begin{bmatrix} u_1^T & u_2^T \end{bmatrix} MM^T \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \sigma^2 \quad \text{and} \quad \begin{bmatrix} -u_2^T & u_1^T \end{bmatrix} MM^T \begin{bmatrix} -u_2 \\ u_1 \end{bmatrix} = \sigma^2 \,,$$

where $u_1, u_2 \in \mathbb{R}^n$, and $\|u_1\|^2 + \|u_2\|^2 = 1$. An analogous relation exists between the eigenvalues and eigenvectors of $A^T A$ and those of $M^T M$. It follows that the singular values and vectors of $A(x)$ are intimately related to those of $M(x)$: All the singular values of A appear twice as singular values of M. In other words, M has the following n pairs of singular values:

$$\begin{aligned} \sigma_1(M) = \sigma_2(M) &= \sigma_1(A)\,, \\ \sigma_3(M) = \sigma_4(M) &= \sigma_2(A)\,, \\ &\vdots \\ \sigma_{2n-1}(M) = \sigma_{2n}(M) &= \sigma_n(A)\,. \end{aligned}$$

Given this correspondence between the singular values of A and those of M, we now shift attention to the smooth (real valued) matrix function $M : \mathbb{R}^3 \to \mathbb{R}^{2n \times 2n}$. Unfortunately, M has many more coalescing points for its singular values than those of A. Namely, as it has just been pointed out, we have:

$$\sigma_{2m-1}(M(x)) = \sigma_{2m}(M(x))\,,$$

for all $m = 1, 2, \ldots, n$ and $x \in \mathbb{R}^3$. But we are really interested only in the coalescing points for the pairs $(\sigma_2(M), \sigma_3(M))$, $(\sigma_4(M), \sigma_5(M))$, etc., as they are the only ones that correspond to coalescing points for A.

2. From the considerations above, we see that the pairs of singular values $(\sigma_1(M)$, $\sigma_2(M))$, $(\sigma_3(M), \sigma_4(M))$, etc., coalesce everywhere in $\mathbb{R}^3$, while the pairs $(\sigma_2(M)$,

$\sigma_3(M))$, $(\sigma_4(M), \sigma_5(M))$, etc., coalesce at isolated points. This is highly non-generic for singular values of a real valued matrix function that depends on three parameters, and is due to the special block-structure of M. But our methods, developed in Chapter 4, are designed to work only in generic situations. So, as it is, the problem of detecting coalescing points for the singular values of M is still unfit to our methods. Now, our idea is to perturb the function M by adding to it a matrix $E \in \mathbb{R}^{2n\times 2n}$, with entries chosen at random and of small norm, that does not have any specific structure. The effect of this perturbation should be to turn our matrix function M into a new one:

$$\widetilde{M}(x) = M(x) + E\,, \forall\ x \in \mathbb{R}^3\,,$$

for which the generic properties of a real valued matrix function that depend on three parameters hold. From now on, we assume that the singular values of $\widetilde{M}$ do coalesce along curves in the three dimensional space, as generically expected.

3. Eventually, we have come to a problem that we can handle: computing the curves along which pairs of singular values of $\widetilde{M}$ coalesce. Now, suppose that the singular values $\sigma_m(A)$ and $\sigma_{m+1}(A)$ coalesce at a point $\xi_0 \in \mathbb{R}^3$. We want to detect such a possibility by looking at the "coalescing curves" for $\widetilde{M}$. We begin by giving a simple perturbative result, adapted from [13].

Theorem 5.2.1. *If M and $M + E$ are in $\mathbb{R}^{2n\times 2n}$, then for $m = 1, \ldots, 2n$:*

$$|\sigma_m(M + E) - \sigma_m(M)| \leq \|E\|\ .$$

This result allows us to make useful considerations about possible location and expected topological properties of some of the coalescing curves in which we are interested. In fact, let

$$\Gamma_{m,m+1} = \{x : \sigma_m(\widetilde{M}(x)) = \sigma_{m+1}(\widetilde{M}(x))\}\,,$$

for $m = 1, \ldots, 2n-1$, and let us focus only on $\Gamma_{2j,2j+1}$, $j = 1, \ldots, n-1$. From the perturbative result above, it follows that, for every $j = 1, \ldots, n-1$, $\Gamma_{2j,2j+1}$ must be

confined within the set

$$\mathcal{S}_j = \{x : |\sigma_{2j}(M(x)) - \sigma_{2j+1}(M(x))| \leq 2\,\|E\|\}\,.$$

In practical cases, if $\|E\|$ is small enough (we have often taken $\|E\| = 10^{-2}$), it is reasonable to expect that each $\mathcal{S}_j$ is the union of (small) closed bounded sets. Moreover, any coalescing point for the pair $(\sigma_{2j}(M), \sigma_{2j+1}(M))$ must also be contained in $\mathcal{S}_j$; therefore, we anticipate branches of the curve $\Gamma_{2j,2j+1}$ to be localized near the coalescing points for the pair of singular values $(\sigma_{2j}(M), \sigma_{2j+1}(M))$, or equivalently for $(\sigma_j(A), \sigma_{j+1}(A))$. Also, the fact that each one of the curves $\Gamma_{2j,2j+1}$ needs to be bounded allows us to anticipate that they will be closed curves.

On the other hand, Theorem 5.2.1 does not allow us to draw any conclusion about the possible location of the curves $\Gamma_{2j-1,2j}$, $j = 1, \ldots, n$, as the corresponding pairs of singular values for M are equal everywhere in $\mathbb{R}^3$.

Comforted by the considerations above, we conjecture that the presence of small (closed) branches of the curve $\Gamma_{2j,2j+1}$, for some $1 \leq j \leq n-1$ will indicate that a coalescing point for the pair $(\sigma_j(A), \sigma_{j+1}(A))$ is close by. This is the backing of our method for the detection of coalescing points for complex valued functions of three parameters. In a nutshell, our idea is to fully describe the set of coalescing curves for $\widetilde{M}$, in a given domain of $\mathbb{R}^3$; once this has been done, use the curves $\Gamma_{2j,2j+1}$ to construct an approximation of the coalescing point for $(\sigma_j(A), \sigma_{j+1}(A))$. This approximation is then corrected via a complex analogous of the zoom-in module of 4.2.

Remark 5.2.2. It is time now to stress the main difference between the "detection" technique we propose for the complex case, and the one used in the real case. Here in the complex case, we use a somewhat empirical approach; the detection, or localization, is based on geometrical evidence, therefore the intervention of a user, capable of interpreting a picture, is essential. On the other hand, the detection approach used

in the real case was more sound, and prone to be readily automated, as it was based on a change of sign.

We conclude this Section with some examples, in which we test the grounding of the arguments described so far, as well as put to work the continuation algorithm presented in Section 5.1.

Example 5.2.3. Let

$$W_1 = \begin{bmatrix} -0.3159\ldots & -0.9480\ldots \\ 0.0562\ldots & -0.5788\ldots \end{bmatrix} + i \begin{bmatrix} 0.0392\ldots & -0.5157\ldots \\ 0.7742\ldots & -0.9266\ldots \end{bmatrix},$$

$$W_2 = \begin{bmatrix} -0.9629\ldots & -0.5787\ldots \\ -0.6047\ldots & 0.8488\ldots \end{bmatrix} + i \begin{bmatrix} 0.1305\ldots & -0.3208\ldots \\ 0.7705\ldots & -0.5499\ldots \end{bmatrix},$$

$$Q = \begin{bmatrix} -0.21132\ldots + i\,0.16300\ldots & 0.52630\ldots + i\,0.80733\ldots \\ -0.49028\ldots + i\,0.82970\ldots & -0.21828\ldots - i\,0.15354\ldots \end{bmatrix},$$

and let

$$A(x_1, x_2, x_3) = Q^T \operatorname{diag}(x_1 + x_2, 2)\, Q + (x_1 - x_2)W_1 + x_3 W_2\,,$$

for all $(x_1, x_2, x_3) \in \mathbb{R}^3$. The entries of the real and imaginary parts of W_1 and W_2 are chosen at random (uniformly distributed in $[-1, 1]$), while Q is a random unitary matrix. The matrix function $A : \mathbb{R}^3 \to \mathbb{C}^{2\times 2}$ has been built so that it is equal to a multiple of the identity matrix for $(x_1, x_2, x_3) = (1, 1, 0)$. Let $\xi_0 = (1, 1, 0)$. By construction, ξ_0 is a coalescing point for the singular values of A. The first thing we do is to explore the region around ξ_0, in search for coalescing curves for the singular values of the matrix function $\widetilde{M}$. In particular, we want to verify the existence of a small closed curve $\Gamma_{2,3}$ in a neighborhood of ξ_0, as it has been predicted above.

In order to start our continuation algorithm, we need an initial point on the curve. We search such point by freezing x_3 to some value close to 0, and swiping the square $\Omega = [0.9, 1.1] \times [0.9, 1.1]$ through our 2-dimensional localization algorithm (see

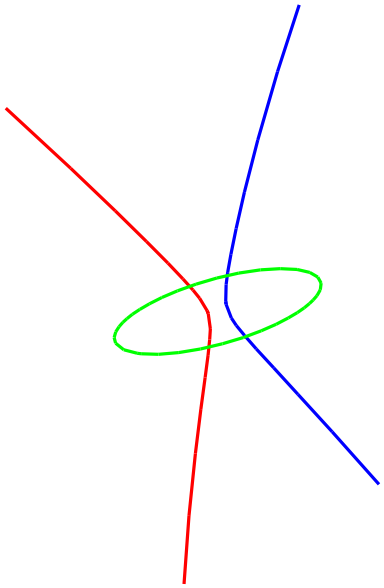

Figure 26: Left (red) curve: $\Gamma_{1,2}$; center (green) ring: $\Gamma_{2,3}$; right (blue) curve: $\Gamma_{3,4}$.

Section 4.1), using a 50×50 grid. After a few attempts, we find that $\widetilde{M}(\cdot, \cdot, 0.01)$ has four coalescing points in Ω, two for the pair (σ_2, σ_3) and one for each of the pairs (σ_1, σ_2) and (σ_3, σ_4). Geometrically, this means that the plane $x_3 = 0.01$ intersects each one of the curves $\Gamma_{1,2}$, $\Gamma_{2,3}$ and $\Gamma_{3,4}$. All four points of intersection are correctly approximated by the zoom-in module. Now we can start the continuation algorithm. We set the absolute value tolerance `tol` in Section 5.1.2 to 10^{-4}, initial step-size 10^{-3}, minimum step-size 10^{-14}, maximum step-size 10^{-1}. Then we launch the continuation algorithm, starting from each of the four points that have been found. The outcome is shown in Figure 26. The picture shows that our insight about $\Gamma_{2,3}$ was indeed correct: The curve is a small loop very close to ξ_0. Indeed, the piercing technique we have adopted has proved effective in detecting that one of the branches is a closed curve.

Remark 5.2.4. Interestingly, we notice that also the curves $\Gamma_{1,2}$ and $\Gamma_{3,4}$ pass very close to the ring $\Gamma_{2,3}$ and ξ_0. This fact has been observed in every experiment that we have conducted. In other words, the configuration shown in Figure 26 has proved to be typical of every coalescing point for the singular values of A. The reason why

Table 5: Performance of the continuation algorithm for the computation of $\Gamma_{2,3}$ in Example 5.2.3.

`tol`	points	average zoom-in iterates	corrector failures	average step-size	cpu-time
10^{-3}	18	5.3	9	4.3×10^{-3}	.37
10^{-4}	39	4.3	–	1.8×10^{-3}	.24
10^{-5}	120	3.7	–	5.5×10^{-4}	.67
10^{-6}	379	3	–	1.7×10^{-4}	1.65

this happens is still not clear to us. Nevertheless, this interesting phenomenon has proved very useful in the detection of coalescing points for the singular values of A, as we will soon appreciate.

Once the loop $\Gamma_{2,3}$ has been detected, we compute its "center" and correct it via the complex analog of the zoom-in algorithm already discussed. In this case, we converge to a point that is within 5.2×10^{-5} of ξ_0, and for which $\sigma_1(A) - \sigma_2(A) \approx 2.1 \times 10^{-11}$.

In Table 5 we give a summary of performance of the continuation algorithm for the computation of the closed curve $\Gamma_{2,3}$ in Figure 26, for several choices of the absolute value tolerance `tol`. Each row of the table shows, in order: Tolerance `tol` used, total number of points computed along the curve, average number of iterates performed by the zoom-in module for each continuation step, number of times the corrector fails to converge, average step-size, total cpu-time required to run the algorithm. We recall that, each time the corrector fails to converge, the step-size is halved and the continuation step retried.

Remark 5.2.5. The numerical experiments of which we have reported in Table 5 have been clocked with the detection feature of Section 5.1.3 turned off, since the use of the this feature heavily penalizes the performance of the continuation algorithm. For example, looking at Table 5, consider the case of `tol` $= 10^{-3}$, for which the continuation without detection feature requires just 0.24 seconds of cpu-time. When the detection feature is turned on, the required cpu-time increases to 3.8 seconds. For

this reason, the use of the detection feature should be limited to situations in which it is strictly necessary.

Example 5.2.6. Let us consider again the function A of the previous example. Note that, at the point $\xi_1 = (-1, -1, 0)$, we have $A(\xi_1) = \mathrm{diag}(-2, 2)$. Therefore, also the point ξ_1 is a coalescing point for the singular values of A. Starting from this consideration, we explore the region around ξ_1 and find, as expected, the "fingerprint" of a coalescing point (see the upper green ring in Figure 27). This time, we decide not to limit ourselves to the region of $\mathbb{R}^3$ close to ξ_1, but to keep exploring the space in search for other coalescing points. To do so, we decide to keep following $\Gamma_{1,2}$ and $\Gamma_{3,4}$, hoping to detect, along the way, the presence of other curves. We set the parameter α (see Section 5.1.3) to 0.1. The continuation of $\Gamma_{3,4}$ doesn't produce anything new. On the other hand, while continuing $\Gamma_{3,4}$, the algorithm detects the presence of other branches of the other curves $\Gamma_{1,2}$ and $\Gamma_{2,3}$. We compute those curves and find another configuration that, in our experience, is clear indication of another coalescing point (see bottom-right green ring on Figure 27) . Finally, we compute the center of the new loop, and zoom in. The zoom-in module converges to a point

$$\xi_2 = (-0.73362\ldots, -2.0624\ldots, 0.18844\ldots)$$

for which $\sigma_1(A) - \sigma_2(A) \approx 7.9 \times 10^{-12}$.

Remark 5.2.7. The way we have detected and approximated ξ_2 is the paradigm of how our technique can be used to find new coalescing points for the singular values of a (three parameter) complex valued function. We summarize it in a more general setting. Given a smooth matrix function $A : \mathbb{R}^3 \to \mathbb{C}^{n\times n}$, we detect, and continue, the curves $\Gamma_{2j-1,2j}$, for $j = 1, \ldots, n$. Those curves do not stay always closed to coalescing points, but rather tend to branch off vast regions of $\mathbb{R}^3$. The technique described in Section 5.2.1 is used to detect the presence of configurations as the one clearly depicted in Figure 26. Those configurations are considered as the "fingerprint" of a

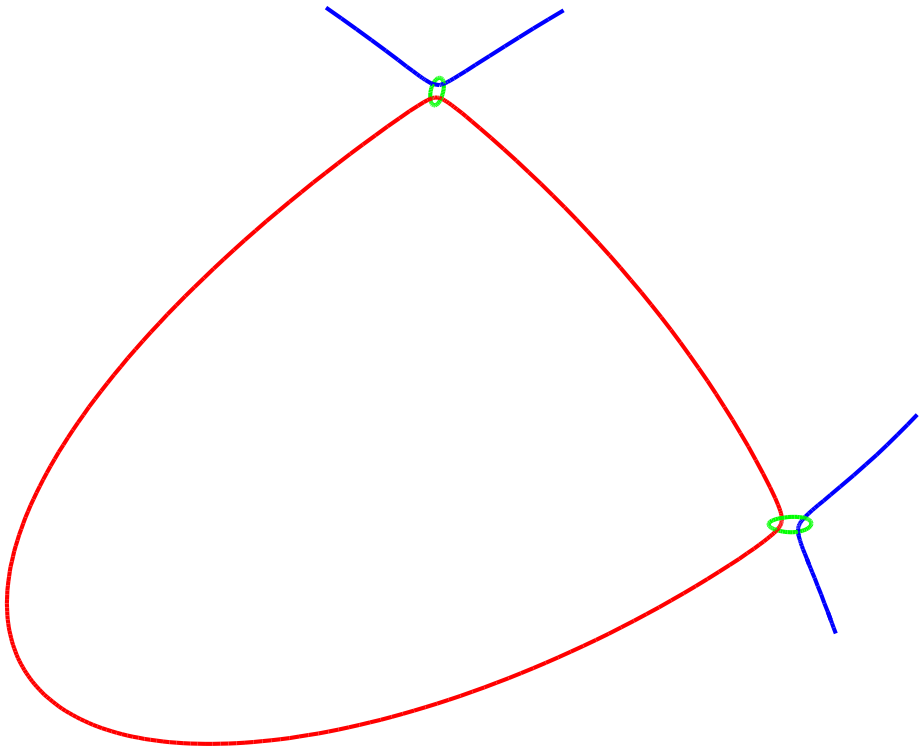

Figure 27: Blue curves: $\Gamma_{1,2}$; green rings: $\Gamma_{2,3}$; red curve: $\Gamma_{3,4}$.

coalescing point. Once such a configuration is found, we use the zoom in module to approximate the coalescing point.

Remark 5.2.8. Our technique for the localization of coalescing points in the complex case is heavily based on the realization that configurations such as the one in Figure 26 reveal the presence of a coalescing point. This is a fact based on numerical evidence, and remains to be proven. Nevertheless, we have verified the effectiveness of our technique, applying it to many different test problems, where (2×2) and (3×3) test matrices have been built similarly to the matrix A of Example 5.2.3.

CHAPTER VI

CONCLUSIONS

In this thesis, we have studied criteria to determine whether a real matrix, smoothly depending on two parameters, has coalescing singular values inside a planar region Ω. The key idea is that, generically, the coalescing of a pair of singular values is related to period-doubling of the corresponding singular vectors, as one completes a loop around the coalescing point. This idea has been used to develop algorithms for the detection and approximation of points where the coalescing occurs. The study has been extended also to the case of real and complex matrices that depend on three parameters.

We conclude with some problems and open questions that we intend to explore in the future:

(i) All the results of Chapter 2, and algorithms of Chapters 4-5, deal with the "full SVD" (or eigendecomposition) of $(n \times n)$ matrices. The computational expenses can be prohibitive if n is large. We intend to adapt our theoretical results and numerical techniques to the case when one is interested only in a small selection of singular values (or eigenvalues); in this case, our algorithms should work with "reduced decompositions" that take into account only the relevant portion of spectrum and invariant subspaces.

(ii) Our localization algorithm of Chapter 4 is very amenable to be parallelized. We intend to exploit this possibility in a future version of the algorithm.

(iii) We have willfully left aside coalescing phenomena that have codimension higher than 2, such as coalescing of two pairs of singular values, triple coalescing of

singular values, and the case $\sigma_{n-1} = \sigma_n = 0$ (see Remarks 1.2.1, 2.1.18, and Example 4.3.5). We intend to explore these cases, as they can occur generically when the number of parameters is sufficiently high.

(iv) The correspondence between the configuration depicted in Figure 26 and coalescing points for complex valued matrix functions, although experimentally evident, still awaits for a rigorous justification.

(v) We intend to test our algorithms on problems that arise in applied sciences, such as Physics, Chemistry and Engineering. Some of these problems (e.g., see the work [26]) appear to posses added structure which lowers the codimension of coalescing phenomena; we are interested in understanding what causes the drop in the codimension and how we can adapt our techniques to these specific cases.

REFERENCES

[1] C. Adams and R. Franzosa. *Introduction to topology pure and applied*. Pearson Prentice Hall, Upper Saddle River, NJ, 2008.

[2] M. Baer. *Beyond Born-Oppenheimer: electronic nonadiabatic coupling terms and conical intersections*. John Wiley & Sons, Hoboken, NJ, 2006.

[3] M.V. Berry, Quantal phase factors accompanying adiabatic changes. *Proc. Roy. Soc. Lond.*, A392:45–57, 1984.

[4] M.V. Berry and M. Wilkinson Diabolical Points in the Spectra of Triangles. *Proc. Roy. Soc. Lond.*, A392:15–43, 1984.

[5] A. Bunse-Gerstner, R. Byers, V. Mehrmann, and N. K. Nichols. Numerical computation of an analytic singular value decomposition of a matrix valued function. *Numer. Math.*, 60:1–40, 1991.

[6] J.L. Chern and L. Dieci. Smoothness and periodicity of some matrix decompositions. *SIAM J. Matrix Anal. Appl.*, 22:772–792, 2001.

[7] L. Dieci and T. Eirola. On smooth decomposition of matrices. *SIAM J. Matrix Anal. Appl.*, 20:800–819, 1999.

[8] L. Dieci, M. G. Gasparo, and A. Papini. Continuation of Singular value Decompositions. *Med. J. of Mathematics*, 2:179–203, 2005.

[9] L. Dieci and A. Pugliese. Two-parameter SVD: Coalescing singular values and periodicity. 2007. Submitted to *SIAM J Matrix Anal. Appl.*.

[10] L. Dieci and A. Pugliese. Singular values of two-parameter matrices: An algorithm to accurately find their intersections. *Computers and Mathematics in Applications*, 2007. To appear.

[11] H. Gingold. A method of global blockdiagonalization for matrix-valued functions. *SIAM J. Math. Anal.*, 9-6:1076–1082, 1978.

[12] H. Gingold and P.F. Hsieh. Globally analytic triangularization of a matrix function. *Linear Algebra Appl.*, 169:75–101, 1992.

[13] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The Johns Hopkins University Press, 2nd edition, 1989.

[14] W. Govaerts. *Numerical methods for bifurcations of dynamical equilibria*. SIAM, New York, 2000.

[15] G. Hernzberg and H.C. Longuet-Higgins. Intersection of potential energy surfaces in polyatomic molecules. *Disc. Faraday Soc.*, 35:77–82, 1963.

[16] M.W. Hirsch. *Differential Topology.* Springer-Verlag, New–York, 1976.

[17] P.F. Hsieh and Y. Sibuya. A global analysis of matrices of functions of several variables. *J. Math. Anal. Appl.*, 14:332–340, 1966.

[18] Tosio Kato. *A short introduction to perturbation theory for linear operators.* Springer-Verlag, New York, 1982.

[19] Keller, H. B. *Lectures on numerical methods in bifurcation problems.* Berlin, Heidelberg, New York: Springer-Verlag. Publ. for TATA Institute of Fundamental Research, 1987.

[20] X. Liu. On a curve veering aberration. *Journal of Applied Mathematics and Physics (ZAMP)*, 25:99–111, 1974.

[21] X. Liu. Behavior of derivatives of eigenvalues and eigenvectors in curve veering and mode localization and their relation to close eigenvalues. *J. Sound and Vibration*, 256:551–564, 2002.

[22] R. Melville and D. S. Mackey. A new algorithm for two-dimensional numerical continuation. *Comput. Math. Appl.*, 30(1):31–46, 1995.

[23] K. A. O'Neil. Critical points of the singular value decomposition. *SIAM J. Matrix Analysis*, 27:459–473, 2005.

[24] M. Ouisse and J.L. Guyader. Localization of structural zones producing hypersensitive behavior: finite element approach. *Comput. Methods Appl. Mech. Engrg.*, 192:5001–5020, 2003.

[25] N.C. Perkins and C.D. Mote Jr. Comments on curve veering in eigenvalue problems. *Journal of Sound and Vibration*, 106:451–463, 1986.

[26] A. Srikantha Phani, J. Woodhouse, and N.A. Fleck. Wave propagation in two-dimensional periodic lattices. *J. Acoust. Soc. Am.*, 119:1995–2005, 2006.

[27] C. Pierre. Mode localization and eigenvalue loci veering henomena in disordered structures. *Journal of Sound and Vibrarion* , 126:485–502, 1988.

[28] R. H. Plaut, K. D. Murphy, and L. N. Virgin. Curve and surface veering for a braced columns. *Journal of Sound and Vibrarion* , 187:879–885, 1995.

[29] M. S. Schuurman and D.R. Yarkony. On the locus of points of conical intersection: Seams near seams. *J. Chem. Phys.*, 126:044104, 2007.

[30] L. Serrano-Andrés and M. Merchán. Computation of conical intersections by using perturbation techniques. *J. Chem. Phys.*, 122:104107, 2005.

[31] Y. Sibuya. Some global properties of matrices of functions of one variable. *Math. Annal.*, 161:67–77, 1965.

[32] J. von Neumann and E. Wigner. Eigenwerte bei adiabatischen prozessen. *Physik Zeitschrift*, 30:467–470, 1929.

[33] David R. Yarkony. Conical intersections: The new conventional wisdom. *J. Phys. Chem. A*, 105:6277–6293, 2001.

[34] David R. Yarkony. On the connectivity of seams of conical intersection: Seam curvature. *J. Chem. Phys.*, 123:204101, 2005.